(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 319 567 B1**

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: **10.02.93** (51) Int. Cl.5: **H04H 9/00**

(21) Application number: **88906360.8**

(22) Date of filing: **10.06.88**

(86) International application number:
PCT/US88/01969

(87) International publication number:
WO 88/10540 (29.12.88 88/28)

(54) BROADCAST INFORMATION CLASSIFICATION SYSTEM AND METHOD.

(30) Priority: **24.06.87 US 66106**

(43) Date of publication of application:
14.06.89 Bulletin 89/24

(45) Publication of the grant of the patent:
10.02.93 Bulletin 93/06

(84) Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

(56) References cited:
EP-A- 0 210 609
US-A- 3 919 479
US-A- 4 230 990
US-A- 4 282 403
US-A- 4 432 096
US-A- 4 450 531
US-A- 4 499 601

(73) Proprietor: **BROADCAST DATA SYSTEMS LIMITED PARTNERSHIP**
1515 Broadway
New York New York 10036(US)

(72) Inventor: **KENYON, Stephen, C.**
230 Tacketts Mill Road
Stafford, VA 22554(US)
Inventor: **SIMKINS, Laura, J.**
12504 Piedmont Road
Clarksburg, MA 20871(US)
Inventor: **SEBASTIAN, Richard, L.**
6128 River Drive
Lorton, VA 22074(US)

(74) Representative: **Melzer, Wolfgang, Dipl.-Ing. et al**
Patentanwälte Mitscherlich, Gunschmann, Körber, Schmidt-Evers, Melzer,Schulz Steinsdorfstrasse 10
W-8000 München 22(DE)

EP 0 319 567 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

Description

This invention relates to apparatus and method by which broadcast information can be recognized and classified. More particularly, this invention relates to a system and method for classifying broadcast information using a plurality of reference signal libraries in a two-stage classification process.

It is known that broadcast stations (television and radio) are regularly monitored to determine when and how often certain information is broadcast. For example, artists may be paid a royalty rate depending upon how often their particular work is broadcast. Likewise, commercial backers of broadcast programming have an interest in determining when and how often commercials are played. Further, marketing executives and the broadcasters themselves are interested in determining the popularity of certain broadcast information in order to target that information to the appropriate audience at the appropriate time. Those of ordinary skill in this field will readily understand that a wide variety of legal, economic and social concerns require the regular monitoring of broadcast information. All such requirements share a common need for certain information such as which information was broadcast and when.

Traditionally, such broadcast station monitoring was performed manually by a plurality of listeners who would physically monitor the broadcast program and manually tabulate which information was broadcast and when. However, the cost of these manual surveys has become prohibitive. Such a method is labor intensive and subject to reliability problems. For example, a manual monitor may easily miss a fifteen second commercial broadcast over radio. In addition, it is virtually impossible for a single individual to monitor a plurality of broadcast channels. Therefore, a great number of monitors has been traditionally required to fully monitor performance in a multi-media environment.

In view of the above problems with manual systems, it has been proposed to design and implement an automatic broadcast recognition system. It is believed that such automatic systems will be less expensive and more reliable than manual surveys.

In recent years, several techniques and systems have been developed which electronically monitor broadcast signals and provide information relative to the content and timing of the program monitored. Initially, these automatic systems performed signal recognition by inserting a code signal in the broadcast signal itself. Upon reception, the automatic system would recognize the code signal (matching it with a reference library) and classify the broadcast information accordingly. Although such coding techniques work for limited applications, they require allocation of portions of the broadcast signal band for identification purposes. In addition, such a system requires special processing, coding and decoding circuitry. Such circuitry is expensive to design and assemble and must be placed at each transmitting and receiving station. In addition, those of skill in this field understand that government regulatory agencies are adverse to providing additional bandwidth for purposes of code signal identification.

To overcome some of the disadvantages involved with the use of the coded signal techniques, certain automatic broadcast signal identification systems have been developed which do not require special coding of the broadcast signal. Such a system is disclosed in U.S. Patent 3,919,479 to Moon et al. In Moon et al, an audio signal is digitally sampled to provide a reference signal segment which is stored in a reference library. Then, when the audio signal is broadcast, successive portions thereof are digitized and compared with the reference segment in the library. The comparison is carried out in a correlation process which produces a correlation function signal. If the reference and broadcast signal segments are not the same, a correlation function with a relatively small amplitude results. On the other hand, if the reference and broadcast signal segments are relatively the same, a large correlation function signal is produced. The amplitude of the correlation function signal is sensed to provide a recognition signal when the amplitude exceeds a predetermined threshold level.

While the Moon et al system may operate effectively in certain situations, it is not effective for many applications. For example, where signal drop-out is experienced, a single segment correlation system may be severely degraded or disabled all together. Additionally, the Moon et al system is relatively insensitive to time-axis variations in the broadcast information itself. For example, it is known that many disc-jockeys "compress" broadcast songs by speeding-up the drive mechanism. It is also known that other disc-jockeys regularly "compress" and/or "stretch" broadcast information to produce certain desired effects in the audience.

In an attempt to overcome such time-axis variations, Moon proposes to reduce the bandwidth of the broadcast signal by envelope-detecting the broadcast signal and providing envelope signals having substantially low, and preferably sub-audio frequency signal components. It has been found that when the envelope signal at sub-audio frequencies is used during the correlation process, the digitally sampled waveforms are less sensitive to time-axis variations. However, the improvements which can be achieved by such a solution are very limited and will only operate for broadcast signals which have been "compressed"

or "stretched" by a small amount. In addition, such a solution is subject to high false alarm rates. These disadvantages make the Moon et al system less than desirable for a rapid, accurate, and inexpensive broadcast information recognition system.

Another automatic signal recognition system is disclosed in U.S. Patent 4,450,531 to Kenyon et al. Mr. Kenyon is a joint inventor of the subject application and the '531 patent. The teachings of the '531 patent are hereby incorporated into this application by reference.

The Kenyon et al system successfully addresses the reliability problems of a single segment correlation system, and the time-axis variation problems experienced by prior systems. In Kenyon et al, a plurality of reference signal segments are extracted from a program unit (song), digitized, Fourier transformed and stored in a reference library in a frequency domain complex spectrum. The received broadcast signal is then prefiltered to select a frequency portion of the audio spectrum that has stable characteristicS for discrimination. After further filtering and conversion to a digital signal, the broadcast signal is Fourier transformed and subjected to a complex multiplication process with reference signal segments to obtain a vector product. The results of the complex multiplication process are then subjected to an inverse Fourier transformation step to obtain a correlation function which has been transformed from the frequency to the time domain. This correlation function is then normalized and the correlation peak for each segment is selected and the peak spacing is compared with segment length. Simultaneously, the RMS power of the segment coincident with the correlation peak segment is sensed to determine the segment power point pattern. Thus, Kenyon et al overcomes the disadvantages of a single segment correlation system by providing a plurality of correlation segments and measuring the distances between correlation peaks. Where the distances match, the broadcast signal is declared as being similar to the signal segments stored in the reference library. In addition, the RMS value comparison operates to confirm the classification carried out using the signal segments.

To overcome the time-axis variation problem, Kenyon et al utilizes an envelope detector and band pass filtering of the broadcast information, similar to the system of Moon et al. In addition, Kenyon et al, proposes the use of more than one sampling rate for the reference signal segments. A fast and slow sample may be stored for each reference signal segment so that broadcast signals from faster rate stations will correlate with the faster rate reference segments and signals from slower rate stations will correlate with the slower rate reference segments. However, the system according to Kenyon et al also suffers from a relatively high false alarm rate and is computationally very demanding. For example, performing the various multi-segment correlations requires a great deal of computer power. Since a multitude of segments are sampled, the system according to Kenyon et al may take a good deal of time and require the use of expensive, powerful computers.

A system for speech pattern recognition is disclosed in U.S. Patent 4,282,403 to Sakoe. Sakoe discloses a speech recognition system in which a time sequence input of pattern feature vectors is inputted into a reference library. The received speech signal is then subjected to spectrum analysis, sampling and digitalization in order to be transformed into a timed sequence of vectors representative of features of the speech sound at respective sampling instances. A time warping function may be used for each reference pattern by the use of feature vector components of a few channels. The time warping function for each reference pattern feature vector is used to correlate the input pattern feature vector and the reference pattern feature vector. The input pattern feature vector sequence is then compared with the reference pattern feature vector sequence, with reference to the warping function, in order to identify the spoken word. However, the Sakoe system time warps the reference patterns rather than the input signal. Thus, a plurality of patterns must be calculated for each reference pattern, necessarily increasing the memory and computational requirements of the system.

A further signal recognition system is disclosed in U.S. Patent 4,432,096 to Bunge. In Bunge, sounds or speech signals are converted into an electrical signal and broken down into several Spectrum components in a filter bank. These components are then integrated over a short period of time to produce the short-time spectrum of a signal. The spectral components of the signal are applied to a number of pattern detectors which supply an output signal only if the short-time spectrum corresponds to the pattern adjusted in the relevant pattern detector. Each pattern detector has two threshold detectors which supply a signal if the applied input lies between the adjustable thresholds. Thus, the pattern detectors supply an output signal only if all threshold value detectors are activated. For each sound of speech, a pattern detector is provided. When a series of sounds is recognized, the series of addresses of the pattern detectors which have successfully generated an output signal are stored and subsequently applied to the computer for comparison. It can be readily appreciated that such a system requires a number of pattern detectors and a corresponding powerful computation device. In addition, while the Bunge system uses a filter bank to provide a low frequency output signal which is relatively less sensitive to time-axis variations, the Bunge

system is still subject to time distortion problems and a high false alarm rate.

Known automatic broadcast recognition systems have been caught in a quandary of choosing an appropriate time-bandwidth (sampling time times frequency band width) product. Where the broadcast signal is sampled with a large time-bandwidth product, signal recognition may be made accurately. However, when a suitably large time-bandwidth product is employed, it will be extremely sensitive to time-axis variations. Thus, most known systems utilize a predetermined time-bandwidth product and suffer recognition inaccuracies and time-axis variations. In addition, the computational load imposed by all known techniques severely limits the number of songs or other recordings that can be simultaneously sampled in real time.

Thus, what is needed is a small, inexpensive system with limited processing power which automatically monitors a plurality of broadcast channels simultaneously for a large number of sounds. Such a system should provide accurate recognition and, at the same time, remain relatively insensitive to time-axis variations.

SUMMARY OF THE INVENTION

The present invention is designed with a view toward overcoming the disadvantages of known automatic broadcast information recognition systems while at the same time satisfying the objectives alluded to above.

These problems and objectives are solved by the method and apparatus for classifying broadcast information comprising the characterising features of independent claims 1 and 14 respectively.

The present inventors have discovered that an inexpensive, reliable and accurate automatic information classification system may be achieved by utilizing a two-stage classification process. First, known broadcast information (a song or commercial) is "played into" the system in order to generate first and second stage reference libraries. Once the libraries have been generated, broadcast information is monitored by the system. In the first stage classification, the input signal is spectrally analyzed and filtered to provide several low bandwidth analog channels. Each of these channels is fed to a feature generator where it is digitized to form a feature data set that is analyzed to determine if it matches one of the patterns in the first stage reference library. In a preferred embodiment the feature generator forms a multi-channel sequence by computing linear combinations of the input channels. Each of these feature sequences is then smoothed using a moving average filter, further reducing the bandwidth. These reduced bandwidth sequences are then resampled to form a feature set of very low bandwidth but long duration. These sequences are grouped into a spectragram and used in the first stage classification process to rule out unlikely candidates in the first stage reference library. In addition, the feature generator generates an additional feature sequence which will be used in the second stage classification process.

Preferably each spectragram is a time/frequency matrix having a plurality of elements. Likewise, the first stage reference patterns are also preferably time/frequency matrices having the same number of elements as the generated spectragram. The first stage classification process then compares the generated spectragram with the first stage reference spectragram. The reference spectragram may be visualized as a template which is "laid-over" the generated spectragram so that corresponding matrix elements match. Then, the difference between corresponding elements of the generated spectragram and the first stage reference spectragram is measured to determine the similarity between the generated spectragram and the reference spectragram. Then, the sum of the element differences for the entire spectragram is obtained to provide a difference measurement between the generated spectragram and the first stage reference spectragram. This difference computation is repeated for each pattern in the first stage reference library. Songs having a difference measurement less than a threshold value are considered to be candidate identifications. Those with difference measurements greater than a threshold value are rejected as being not similar to the broadcast information.

The first stage reference patterns which have not been rejected in the first stage classification process are then queued according to their difference measurements. Thus, the queueing order places the most similar first stage reference pattern at the head of the queue.

Next, the second stage reference classification process is carried out in the queueing order established in the first stage classification process. The second stage reference library contains a number of signal patterns which corresponds 1-to-1 with the entries of the first stage reference library. The second stage reference patterns are queued in the order established in the first stage classification process, and then correlated with the additional feature sequence provided from the feature generator. This additional feature sequence does not have as low a bandwidth as the feature sequence used in the first stage classification process. In a preferred embodiment, a cross correlation is conducted between the additional feature sequence and the second stage reference patterns in the queueing order. If the peak correlation value for

any of the cross-correlations exceeds a detection threshold, the broadcast information is classified as being similar to the information represented by the second stage reference pattern. At this time a recognition may be declared and the time, date and broadcast information identification, and broadcast information source may be entered in a detection log.

By performing the computationally demanding cross-correlation in the queueing order established in the less-demanding first stage classification process, processing resources are conserved and the computer power required to classify broadcast information is significantly reduced.

To account for time-axis variations in the broadcast information, a preferred embodiment may include a "time-warping" function for use in the second stage classification process. Specifically, the additional feature sequence provided to the second stage classification process may be "compressed" and/or "stretched to account for variations in broadcast speed. Then, the second stage correlation process correlates the second stage reference pattern with the unmodified additional feature sequence, with a "compressed" additional feature sequence, and/or with a "stretched" additional feature sequence. Therefore, proper identification can take place even if the broadcast information is broadcast more rapidly or more slowly than intended.

BRIEF DESCRIPTION OF TEE DRAWINGS

The advantageous features according to the present invention will be readily understood from the description of the presently preferred exemplary embodiment when taken together with the attached drawings in which:

FIG. 1 is a block diagram depicting the system according to the presently preferred embodiment;

FIG. 2 is a block diagram showing the filter banks of FIG. 1;

FIG. 3 depicts a series of waveforms showing the wave-shaping carried out in the filter banks of FIG. 2;

FIG. 4 is a series of waveforms showing four feature sequences generated by the feature generator processor of FIG. 2;

FIG. 5 depicts a spectragram which is constructed with relation to the waveforms shown in FIG. 4;

FIGs. 6(a) and 6(b) depict the first stage comparison process carried out between the generated spectragram and the first stage reference matrix;

FIG. 7 shows time-warped versions of the input waveform used in the second stage classification process;

FIG. 8 is a series of waveforms depicting the digitized feature sequence used in the second stage classification process, the normalized second stage reference pattern, and the cross-correlation function therebetween;

FIG. 9 is a top-level flow chart depicting a method according to the preferred embodiment;

FIG. 10 is a flow chart of one step of FIG. 9;

FIG. 11 is a flow chart of another step of FIG. 9;

FIG. 12 is a flow chart of yet another step from FIG. 9;

FIG. 13 is a flow chart showing still another step from FIG. 9; and

FIG. 14 is a flow chart showing the confirming step according to FIG. 9.

DETAILED DESCRIPTION OF THE PRESENTLY PREFERRED EXEMPLARY EMBODIMENT

While the present invention will be described with reference to a broadcast music classification system, those with skill in this field will appreciate that the teachings of this invention may be utilized in a wide variety of signal recognition environments. For example, the present invention may be utilized with radio, television, data transfer and other broadcast systems. Therefore, the appended claims are to be interpreted as covering all such equivalent signal recognition systems.

First, an overview of the invention will be provided for clarification purposes. Reference may be had to FIG. 9 for this overview. The automatic recognition of broadcast recordings is useful in determining the rate of play and the time of play of these recordings to determine royalties, projected sales, etc. The prior art in this area has met with limited success due to the necessity of using a relatively large bandwidth product to ensure accuracy. When a suitably large time-bandwidth product is employed, most techniques experience difficulty due to speed variations common in broadcast music stations. In addition, the music computation load imposed by these techniques limits the number of songs or other recordings that can be simultaneously searched for in real time.

The present invention manages the large processing load imposed by a large signature data base through the use of an efficient, though less accurate, first (screening) stage classification process. This first

stage eliminates from further consideration songs that are clearly different from the current input signal. Only those patterns that match the input reasonably well in the first stage are queued for intensive scrutiny by the accurate but computationally demanding second stage. This is because the queueing order established in the first stage classification process is ranked by order of similarity to the broadcast song. Thus, the second stage will first consider the most likely candidates before the less likely candidates. Early recognition will result with a corresponding decrease in computer processing power. This two-stage classification process results in a classification system whose overall capacity is increased by over an order of magnitude compared to known classification systems.

The use of additional stages in a hierarchical classification structure could provide additional capacity for a given processing resource. For example, much larger data bases may require a three-stage classification process to again conserve processing power. Those of skill in this field will readily understand that the teachings in this invention may also be applicable to a three-or-more stage classification process.

FIGs. 1 and 9 depict apparatus and method according to the presently preferred embodiment. The audio signal from one or more broadcase sources is input to an audio channel receiver 4 through antenna means 2. In the present invention, the processing structure of FIG. 1 allows simultaneous processing of up to five audio channels. Therefore, up to five broadcast stations may be monitored and their broadcast programs classified. Additional hardware and software modifications could be performed to increase or decrease the number of channels simultaneously monitored.

From audio channel receiver 4, the input audio signal is provided to an audio preprocessor 6. Audio preprocessor 6 may include filter banks 8, envelope detector 10, and low pass filters 12. The audio preprocessor performs a coarse spectral analysis by bandpass filtering the audio input into several bands. These bands are then envelope detected and lowpass filtered to form several low bandwidth analog channels. Each of these channels is then fed to a processor which performs a feature generation function. Specifically, the processor digitizes and further processes the low bandwidth analog channels to form a feature data set that is analyzed to determine whether it matches one of the patterns of the first stage classification library. In the preferred embodiment, the feature generating processor forms a four-channel sequence by computing linear combinations of the input channels. Each of these feature sequences is then smoothed using an averaging filter, further reducing the bandwidth. These reduced bandwidth sequences are then resampled to form a feature set of very low bandwidth but long duration.

The sequences are then grouped into a spectragram and used by the first stage classification process to rule out unlikely candidates. In addition, a fifth sequence is generated by the feature generating processor for use in the second stage classification process.

The four input channels are combined in a weighted sum to form a feature sequence with specific properties. In the process of forming the linear combinations, weighting coefficients are used which are specially selected to minimize the influence of broadband impulsive energy. It has been found that this greatly reduces sensitivity to speed variations and ampiitude distortions that frequently result from the use of compressors by broadcast stations.

The second stage feature sequence provided by the feature generating processor is not filtered and resampled, but is used at a relatively large bandwidth. Since it is used at this greater bandwidth, a feature sequence that is long enough to provide satisfactory pattern discrimination will be very sensitive to speed variations. To counter this, the sequence is resampled at slightly different rates to form several new sequences that represent the input waveform at different speeds. This process is referred hereinafter as "time warping". A recording that is broadcast faster than normal must be expanded or stretched to replicate the original waveform. Similarly, a recording that is broadcast slower than normal must be compressed. The set of compressed and expanded waveforms comprise a linearly time warped set of replicas of the input feature sequence.

The first classification stage operates on the low bandwidth spectragrams, treating them as time-frequency matrices of the most recent feature data. These spectragrams should be normalized to compensate for gain variations in the broadcast and receive systems.

In the first stage classification process, the most recently generated spectragram is compared with every reference pattern stored in the first stage reference library. For each first stage reference pattern in the library, a distance is computed between it and the input spectragram. This distance may be computed as the sum of differences between corresponding elements of the input and reference matrices. This distance is a measure of the similarity between the current broadcast song and the subject reference recording. This distance computation is repeated for each pattern in the data base. Songs whose distances are less than a threshold value are considered to be candidate identifications. Those with distances greater than the threshold value are rejected.

To conserve processing resources and insure that the most likely candidates are considered first, the

first stage classifications are positioned in a queue according to their distance scores. Patterns that have passed the first stage classification test and entered into the queue are subject to a confirming classification in the second stage. This classification process uses the single channel wider bandwidth feature set, including the time warped replicas. For each entry in the queue, the corresponding single channel reference pattern is compared with each of the time warped replicas of the most recent feature vector. A correlation procedure is employed that involves computing the cross correlation function and then scanning it to select the maximum value. This is repeated for each of the time warped replicas. If the peak correlation value for any of the correlations exceeds a detection threshold, a recognition is declared and the time, date, song identification number, and radio station are entered in a detection log. If none of the songs in the queue passes the confirming classification, the next time segment is analyzed in the same way. In such a fashion, an inexpensive, efficient and accurate broadcast music classification system is realized. Therefore, a small system with limited processing power can monitor several radio channels simultaneously for a large number of songs. This large capacity has an economic advantage in that the revenue producing capability of a single monitoring unit is proportional to the number of songs monitored times the number of stations under surveillance.

The first stage features have a reduced dimensionality and can be computed and evaluated at low computation cost. Nevertheless, the effect of the first stage processing is to reduce the number of song candidates to a small fraction of the data base size. The second stage of classification uses a song signature of significantly higher dimensionality (time band width product) than the first stage, makes song detection decisions only for songs queued in the first stage, and only for those songs which are identified as probable for the first stage. The first stage has a song detection threshold bias towards high detection rates and moderate false alarm rates. The second stage has both high detection rates and low false alarm rates. The net effect of this two stage detection procedure is the ability to monitor a large number of songs over several channels using only limited processing power. Thus, the apparatus and method according to the present invention may provide an economically significant broadcast information classification system.

Turning now to FIG. 1, the apparatus according to the present invention will be described. Antenna 2 receives radio waves including audio signals. The antenna apparatus is capable of receiving up to five radio channels simultaneously. The audio signal is received by audio channel receiver 4, and provided to audio preprocessor 6 through MULTIBUS (TM of Intel Corp.) 100. Audio preprocessor 6 includes filter banks 8, envelope detectors 10, and low pass filters 12. The audio preprocessor 6 will be described in more detail with reference to FIG. 2.

FIG. 1 also depicts analog-to-digital converter 14 which may be used to digitize the audio signal. Multiplexer 16 is used to carry out multiplexing operations when a plurality of audio signals is being simultaneously classified. Both A/D converter 14 and multiplexer 16 are also coupled to MULTIBUS (TM of Intel Corp.) 100.

Also coupled to MULTIBUS (TM of Intel Corp.) 100 is an array processor 18. Array processor 18 comprises a Marinco 3024 CPU 20 and a feature vector operations section 22. Both the feature vector operations section 22 and the CPU 20 are coupled to MULTIBUS (TM of Intel Corp.) 100. The functions of array processor 18 include the time warping of the second stage feature sequence and the second stage correlation computations.

The processor 24 is also coupled to MULTIBUS (TM of Intel Corp.) 100 and performs the functions of control, data-base management, all in/out (I/O) management, and the first stage classification calculations. Processor 24 may include a microprocessor 26, a memory 28, I/O interfaces 30, a real time clock 32, reference pattern memory 34 and off-line memory 36. Preferably, microprocessor 26 may be a Motorola 68000 series microprocessor. Preferably, working memory 28 includes two Megabytes of memory. Likewise pattern memory 34 stores both the first stage and second stage reference libraries and preferably is realized by a 10 Megabyte hard disk. The off-line memory 36 may be used to change/add/delete reference patterns from the reference pattern libraries in memory 34. Preferably off-line memory 36 comprises a one Megabit floppy disk.

Finally, the processing system may be coupled with such peripherals as CRT 38, printer 40, and terminal 42. Such peripherals are coupled to the system through I/O interfaces 30.

Turning now to FIG. 2, the coarse spectral analysis step S100 of FIG. 9 will be described. The received audio signal is provided to audio preprocessor 6 where it is divided into a plurality of channels. In the presently preferred embodiment, four channels have been selected. However, greater or fewer channels may be used depending upon the exact type of information which is to be classified. Each channel includes a bandpass filter 8, a rectifier 10, and a low pass filter 12. The purpose of the audio preprocessor is to reduce the amount of information processed in the first stage. This provides a long term averaging of the first stage features. Since the purpose of the first stage is to reduce the computation required for

recognition, it is desirable to reduce the amount of information processed per unit time. Signal discrimination accuracy is proportional to the time bandwidth product of the feature vector. Therefore, by reducing feature vector bandwidth while expanding duration, accuracy is maintained while required processing per unit time is decreased. This is true for any process that requires continuous searching for time series events.

In order to accomplish this, the audio input signal depicted in FIG. 3 is provided to each of bandpass filters 8. Each bandpass filter outputs a signal depicted in FIG. 3 as the bandpass filtered signal. The filtered signals are provided to rectifiers 10, each of which outputs a waveform shown in FIG. 3. Finally, the rectified signals are provided to lowpass filters 12, each of which outputs a lowpass filtered signal, as depicted in FIG. 3. By sampling the reduced bandwidth signal, processing time is conserved while simultaneously reducing the sensitivity of the system to speed variations in the audio signal. Therefore, from lowpass filters 12 are provided a plurality of waveforms as depicted in FIG. 4. These waveforms are respectively denoted by $X_1(t)$, $X_2(t)$, $X_3(t)$, and $X_4(t)$. Each of these waveforms is provided to processor 24 which generates feature sequences according to the waveforms.

Processor 24 thus provides a plurality of feature sequences denoted by $X_{S1}(t)$, $X_{S2}(t)$, $X_{S3}(t)$, $X_{S4}(t)$ and $X_c(t)$. Each of these feature sequences 25 is formed as a linear combination of all four waveforms $X_1(t)$ through $X_4(t)$. As shown in FIG. 4, at a time t_1 the four waveforms are sampled and amplitude voltages V_{A1}, V_{B1}, V_{C1} and V_{D1} are respectively measured. Then, for time t_1 a feature vector may be calculated for each of the waveforms. The feature vector is a series of numbers describing characteristics of the input signal. In the preferred embodiment, a feature vector for waveform $X_{S1}(A)$ at time t_1 may be calculated as follows:

$$X_{S1}(t_1) = K_1V_{A1} + K_2V_{B1} + K_3V_{C1} + K_4V_{D1} \quad (1)$$

Thus, each sequence of feature vectors includes components from each of the four waveform bands. The coefficients K are selected to specifically suppress noise.

The special selection of coefficients K is used to suppress the effects of amplitude distortion in the broadcast signal. Amplitude distortion (sometimes denoted as "amplitude compression") sometimes is intentionally applied by certain broadcast stations to avoid overdriving inexpensive receivers. Such "amplitude compression" degrades the similarity of a stored reference pattern to that computed from the input radio signal. For a given level of detection of reliability, this requires larger reference patterns to be used than would be necessary if the distortion were not present. The need for large reference patterns causes a reduction in processing efficiency, particularly making it difficult to employ an effective first stage which makes preliminary decisions using song signatures of low dimensionality. The newly developed approach overcomes this distortion problem by taking explicit advantage of the spectral Properties of this distortion.

The "amplitude compression" process does not significantly effect narrowband signal components, but primarily affects impulsive components which are wideband in nature. A multi-channel time series consisting of a frequency based time series before compression will be denoted as $f_i(t_j)$. After compression, each band time series becomes:

$$g_i(t_j) = a_i(t_j) + f_i(t_j) \quad (2)$$

where compression is described as an additive component $a_i(t_j)$ to each band. In the implemented approach to suppressing effects of amplitude compression, the additive components $a_i(t_j)$ are assumed to be linearly related. Thus, it is assumed that there exists a linear equation which approximately estimates each $a_i(t_j)$ based on the values of $a_k(t_j)$ for $k \neq i$. Thus,

$$a_i(t_j) = \sum_{k \neq i} b_{ik}a_k(t_j) + e_i(t_j) \quad \ldots (3)$$

where $e_i(t_j)$ is the estimation residual for time t_j.

If the f_i may be approximated as independent (which they are not), then the coefficients b_{ik} may be estimated by the correlation process from a data epoch covering time (t_1, t_M) through the solution of the set of (N-1) linear equations:

$$(a_i, a_k) - \sum_{k \neq i} b_{ik'} (a_{k'}, a_k) = 0 \qquad \ldots (4)$$

for each $k \neq i$ where

$$(a_k, a_{k'}) = \sum_{j=1}^{M} a_k(t_j) a_{k'}(t_j) \qquad \ldots (5)$$

Since the $[f_k]$ are approximate as independent, it follows that $(a_k, a_{k'}) \cong (g_k, g_{k'})$ for $k' \neq k$.

This estimate of $(a_k, a_{k'})$ is most accurate when the $a_k(t_j)$ take on their largest magnitude in comparison with the $f_k(t_j)$. This occurs when amplitude compression is the greatest. Thus, the estimate is made:

$$(a_k, a_{k'}) \cong (g_k, g_{k'}) \qquad (6)$$

where $(g_k, g_{k'})' = \Sigma\, g_k(t_j) g_{k'}(t_j)$ (7)

$\{j : |g(t_j)| > T\}$

where $|g(t_j)|$ is a measure of the magnitude of the received broadcast signal and T is a selected threshold above which the signal is considered to be heavily compressed.

The set of linear equations is solved for estimates of the b_{ik}. Then, the effect of compression is suppressed by replacing received energy band time series $g_i(t_j)$ by:

$$g'_i(t_j) = g_i(t_j) - \sum_{k \neq i} b_{ik}\, g_k(t_j) \qquad \ldots (8)$$

To a linear approximation, the N time series $g'_i(t_j)$ have had the effect of compression removed. This approach suppresses linearly dependent information between energy bands and emphasizes linearly independent information. The linearly dependent information can be added, for improved recognition purposes, but must be downweighted because of its vulnerability to amplitude compression.

What is achieved by this method is a set of $g'_i(tj)$ which are relatively immune to the effects of compression. The coefficients b_{ik} may be estimated using data from an ensemble of broadcast music and transmitting stations so that they do not have to be re-estimated for each broadcast station, and so that they are independent of the music being transmitted.

Sampling of the waveforms in FIGURE 4 is preferably conducted at a rate of eight times per second. The bandwidth is preferably $1\frac{1}{2}$ Hz. The Nyquist sampling theory indicates that the sampling rate should be approximately 4 Hz. The present inventors have chose a sampling rate of 8 Hz in order to ensure greater accuracy.

Referring now to FIGURE 5, it can be seen that processor 24 constructs a spectragram in accordance with the linear combinations of the waveforms of FIGURE 4. Thus, as shown in FIGURE 5, each block contains data integrated over eight seconds of time. Thus, the spectragram is a matrix having four spectral channels and eight time channels. Each matrix element contains a feature component calculated as described above. The spectragram is computed as indicated in step S110 of FIGURE 9. This will be described in more detail with reference to FIGURE 10.

According to FIGURE 10, the feature data sets determined in step S100 are smoothed using a moving average filter, as indicated in step S101. Next, at step S103, the waveforms are resampled to form a low bandwidth multi-channel time series. This has the further advantage of reducing the sensitivity to speed variations. Finally, at step S105 the time/frequency matrix of the most recent sample is formed, as depicted in FIGURE 5.

Once the first stage spectragram has been generated, the spectragrams should be normalized to compensate for gain variations in the broadcast and receive systems. This step is depicted as S130 in FIGURE 9, and is further illustrated in FIGURE 11. To accomplish the normalization, all elements of the input spectragram are summed in step S111. This sum represents the total energy in the spectragram.

Then, at step S113, each element of the spectragram is divided by the spectragram sum. This produces a spectragram with unit energy for easy comparison with the reference pattern.

After normalization, the input spectragram and the reference spectragram from the first stage reference library are compared in a preliminary classification step S150, as shown in FIGURES 9 and 13. Each element of the input and reference spectragrams preferably includes 16 bits representing the value of the matrix element. As visually depicted in FIGURES 6(a) and 6(b), the first stage comparison is merely a matter of matching the input signal spectragram with the reference matrix. This process can be visualized as overlaying a reference template on the input signal spectragram.

Since each of the signal and reference matrices contains the same number of elements, a 1-to-1 comparison between matrix elements is conducted. As shown in FIGURES 6(a) and 6(b), the value of matrix element $X_{S1,1}$ is compared with matrix element $X_{R1,1}$. This may be visualized as comparing the distances between the two element values. Returning to FIGURE 12, the distance between the input spectragram is determined in step S131. This is accomplished by summing the differences between corresponding time/frequency elements of the signal spectragram and the reference matrix. This distance calculation is carried out for each-of the entries in the first stage reference library.

Next, at step S133, all first stage reference matrices whose distance measurements are less than a predetermined threshold are accepted as likely candidates. Those first stage reference matrices whose distance measurements exceed the threshold are rejected as unlikely candidates.

Once a distance measurement has been calculated for each matrix in the first stage reference library, those songs that are identified as likely candidates are subjected to a sort and queueing step S170, as depicted in FIGURES 9 and 13. As discussed above, by queueing the songs in their order of similarity to the input signal, the computationally demanding second stage classification will be greatly abbreviated. It should be noted that a wide variety of sort and queueing procedures are available for carrying out this step. The inventors have decided to utilize the sorting and queueing procedure depicted in FIGURE 13.

At step S151 of FIGURE 13, the distance value for each queue entry is set to a maximum. Next, for each song whose distance measurement is less than the threshold value a queue entry is generated containing the song number and its distance score, as shown in step S153. Then, for each new entry into the queue, the queue is scanned from the end to locate the rank order position for the new entry. The new entry is then inserted into the queue at the appropriate space. Entries having a larger distance than the new entry will then be moved toward the end of the queue. This process is depicted in step S155. Lastly, in step 157, the array processor is directed to process songs in the queue in ascending order of distance measurements. Thus, a reference spectragram which has a low distance value from the input song will be subjected to the second stage classification before a reference spectragram having a higher distance measurement.

The above-described procedures complete the first stage classification process. The first stage produces a queue ordered by similarity to the input song. This queue order will be used in the second stage classification process to compare the most likely reference songs to the input song.

Second stage classification actually begins with the generation of the wider bandwidth feature sequence $X_c(t)$, as depicted in FIGURE 2. As discussed above, it is necessary to "time warp" the second stage feature sequence in order to account for speed variations in the broadcast signal. For purposes of the preferred embodiment, it is assumed that all such speed variations are constant, and thus the time warping feature of the preferred embodiment is a linear time warp.

Radio stations are known to introduce significant speed variations into recordings that are broadcast. For a feature vector with a sufficient time-bandwidth product to provide near error free recognition, most recognition systems are intolerant of these speed variations. This problem has been addressed in the Kenyon et al patent referred to above through the use of a segmented correlation approach. In this approach, short feature vectors with relatively small time-bandwidth products were identified separately. The final recognition decision was based on the timing of individual segment recognitions. While this procedure allowed recognition of songs with substantial speed variation, it did not take full advantage of the fact that these speed differences introduce a linear time base error. The method according to the present invention is to linearly compress or expand the time base of the input feature sequence until the distortion has been removed. The entire feature vector is then compared with undistorted vectors from the reference set. The compression and expansion of the feature vector is performed by resampling and interpolating the input sequence. This linear time warping is done in small enough increments that at least one increment will match the stored reference with essentially no degradation. Such time warping is depicted in FIGURE 7.

As shown in FIGURE 7, the input waveform can be compressed into a compressed waveform, and/or stretched into an expanded waveform. According to the preferred embodiment, a set of four time warped waveforms are provided in addition to the un-warped waveform. For purposes of broadcast music

recognition, applicants have chosen to provide waveforms compressed by 2% and 4%, and waveforms stretched by 2% and 4%. Thus, two compressed waveforms, two stretched waveforms, and the original waveform are provided for comparison to the second stage reference library.

Next, as depicted by step S270 in FIGURE 9, a confirming classification is carried out between the time warped (and un-warped) waveforms and the reference patterns in the second stage reference library according to the queueing order established in step S170.

Generally, this confirmation classification is carried out in accordance with the teachings of the Kenyon et al patent incorporated herein by reference. In brief, a correlation is a mathematical means of measuring how well two waveforms compare. If one of the two waveforms is of finite length and is permanently stored in an appropriate memory storage device, a running comparison of the finite stored waveform against a second continuous waveform may be accomplished through on-line solution of the correlation integral equation to produce a third waveform known as the correlation function. When the continuous waveform contains a signal segment (which may even be obscured by noise) which matches the stored waveform, the correlation function attains a large value. The sensing of this large value constitutes a recognition, and is the process by which the occurrence of a commercial advertisement or song is recognized in the process according to the present invention.

The pattern matching according to the present invention involves a correlation procedure with a minimum of 50% overlap of input data and zero filling of reference patterns so that linear correlations are generated instead of circular correlations. The correlation procedure involves computing the cross correlation function and then scanning it to select the maximum value. This is repeated for each of the time warped replicas. If the peak correlation value for any of the correlations exceeds a detection threshold, a recognition is declared and the classification process for that song is over. The confirmatory classification process will now be described with reference to FIGURES 8 and 13.

In FIGURE 8, the digitized broadcast waveform may be one of the time warped (or un-warped) waveforms generated in step S210. For example, this digitized broadcast may represent 512 samples of the audio signal input, taken over a 64 second period. Next, a normalized reference pattern from a second stage reference library is matched to an arbitrary portion on the digitized broadcast waveform. Note that only the first half of the normalized reference contains a waveform, the second half being zero filled. Zero filling is used to accomplish segment splicing which takes the newest reference block and concatenates it to the previous block. Next, a cross correlation is carried out between the digitized broadcast and the normalized reference to provide a correlation function, as depicted in FIGURE 8. Note that correlation peak CP indicates a high correlation between the digitized broadcast and the normalized reference at that particular point.

The correlation process is carried out by first computing the Fourier transform of all five time warped and un-warped waveforms. This provides complex conjugate spectra which are compared with the second stage reference patterns. The reference patterns themselves have been previously normalized so that no additional normalization is required for the reference patterns. Next, samples from the digitized broadcast and reference waveform are cross multiplied and inverse Fourier transformed to provide the correlation signal depicted in FIGURE 8. Note that the correlation function in the zero filled half of the normalized reference waveform is minimal. Thus, only correlations in the first half of the correlation function are valid. The second half of the correlation function is generated by taking the correlation waveform and essentially reversing it to get a mirror image of the reference waveform. The above-described cross correlation process is depicted at step S211 in FIGURE 14.

Next, the correlation functions between each second stage reference pattern and the plurality of time-warped (and un-warped) input signals are compared to select the maximum correlation value for the current input song, as depicted in step S213. The appropriate waveform with the highest correlation value is selected and compared to a threshold value which determines recognition, as depicted in step S215. As soon as a correlation peak value is determined to be above the pre-set threshold, a victory is declared and it is determined that the song has been "recognized". Then, the time of detection, the date, the broadcasting station, and the song number may be derived and provided to output peripheral equipment. Such decision logging may be carried out as step S300, depicted in FIGURE 9. This completes the second stage classification process.

Thus, the above-described system and method provides for an accurate, reliable, yet inexpensive system for classifying broadcast information. Those of skill in this field will understand that a wide variety of modifications may be made without departing from the spirit and scope of the subject invention.

An additional advantage of the apparatus according to the present invention is that it may be used to generate the first and second stage reference libraries. This is an automatic training procedure in which broadcast songs are "played into" the system to provide the first and second stage reference patterns. The

automatic training procedure first selects the most spectrally distinctive section of the song for use in reference pattern generation. Certain sections of a recording are more spectrally distinctive than others. When such a section is used to generate a reference pattern, the performance of a pattern recognizer is improved since it operates on more information. A measure of the distinctiveness of a portion of a recording is the bandwidth of this feature vector. This can be estimated as follows:

$$B=[\sum_{0}^{W} X(w)]^2/[\sum_{0}^{W} X^2(w)] \quad \ldots (10)$$

where X(w) represents the power spectral density at any particular frequency. Very large bandwidths can be produced by songs with impulsive features. While these features are distinctive, they are more subject to distortion and require greater processor dynamic range.

Areas containing impulsive features may be located by computing the crest factor as a ratio of the peak feature value to the standard deviation of the feature vector computed in the same region. A composite figure of merit is then computed as either the ratio of, or the difference between, the bandwidth and crest factor of the second stage feature vector. This is repeated in small increments of time (for example one second) throughout the song. At positions where the figure of merit is highest, a first stage reference feature matrix is computed and tested for time alignment sensitivity. For timing errors as large as can be encountered due to offsets and time scale errors, the resultant distance must remain below a threshold value. The position in the song for training is selected as the one with the highest second stage figure of merit that also passes the first stage time sensitivity test. Those of skill in this field will appreciate that the same hardware used to conduct music classification can also be used to generate the first and second stage reference libraries.

Therefore, what has been described above is apparatus and method for automatically classifying broadcast information according to stored reference patterns. Since the system is microprocessor based, it can be realized in an extremely small and economical package. The costs of constructing and installing such systems will be economically advantageous. Those of skill in this field will readily understand the advantages achieved by the structure and functions of the above-described invention.

While the present invention has been described in connection with what is presently considered to be the most practical and preferred embodiments, it is to be understood that the invention is not limited to the disclosed embodiment.

Claims

1. Method of classifying broadcast information, comprising the steps:
   receiving broadcast information;
   correlating (S270) the broadcast information with a library of second stage reference pattern; and
   classifying (S270, S300) said broadcast information as similar to one of said second stage reference patterns based on said correlating step,
   characterized in that
   between receiving and correlating said broadcast information is compared (S150) with a library of first stage reference patterns, said first stage reference patterns being queued (S170) in an order of their similarity to said broadcast information; and
   that said library of second stage reference patterns corresponds to said first stage reference patterns in the queueing order established in the queueing step.

2. A method according to claim 1 further including the steps of:
   generating a plurality of analyzed waveforms from said broadcast information;
   time warping one of said analyzed waveforms to provide at least one time warped waveform; and
   wherein said correlating step includes the step of correlating both said one analyzed waveform and said time warped waveform with said library of second stage reference patterns.

3. A method according to claim 2 wherein said time warping step includes the step of linearly time warping said one analyzed waveform to provide a stretched waveform.

4. A method according to claims 1 wherein said receiving step includes the step of simultaneously receiving a plurality of broadcast information, and wherein said steps of comparing, queueing, correlating, and classifying are performed on said plurality of broadcast information, substantially simultaneously.

5. A method according to claim 1 further including.the steps of:
processing said broadcast information to provide a plurality N of analyzed signal patterns corresponding to said broadcast information;
generating a spectragram from said analyzed signal patterns;
providing said spectragram to said comparing step for comparison to said first stage reference patterns; and
providing one of said analyzed signal patterns to said correlating step for correlation with said second stage reference patterns.

6. A method according to claim 5 wherein said step of processing includes the steps of:
bandpass filtering said broadcast information into a plurality of bands; and
computing a plurality of linear combinations of said plurality of bands to provide said plurality of analyzed signal patterns.

7. A method according to claim 6 wherein said step of bandpass filtering includes the steps of bandpass filtering, rectifying, and then lowpass filtering each said band.

8. A method according to claim 5 wherein said step of generating a spectragram includes the steps of:
sampling said plurality of analyzed signal patterns at a predetermined rate;
constructing a time/frequency matrix having N frequency channels, n time channels, and a plurality of N(n) of matrix elements;
calculating, for each matrix element, a matrix value $X_{N,n}$ as follows:

$$X_{N,n}(t) = [K_1 \times V_{An}(t)] + \ldots [K_N \times V_{Nn}(t)]$$

where: t equals time at which samples are taken from said plurality of analyzed signal patterns;
V_{An} through V_{Nn} are amplitude values of first through N^{th} analyzed signal patterns taken at sample time t; and
K_1 through K_N are constants preselected to minimize the influence of broadband impulsive energy.

9. A method according to claim 8 further including the step of normalizing said matrix to provide said spectragram.

10. A method according to claim 8 wherein each said first stage reference pattern comprises an Nxn reference matrix of elements, and wherein said step of comparing comprises the steps of:
measuring a variation between each element of the time/frequency matrix and a corresponding element of each of said reference matrices;
summing the measured variations between the time/frequency matrix and each reference matrix to provide a distance measurement for each reference matrix;
comparing each distance measurement with a threshold value; and
eliminating those first stage reference patterns whose corresponding distance measurement exceeds said threshold value.

11. A method according to claim 10 wherein said queueing step comprises the step of ordering non-eliminated first stage reference patterns according to their corresponding distance measurements.

12. A method according to claim 1 wherein said correlating step includes the steps of:
calculating a correlation value for at least one second stage reference pattern with reference to said broadcast information; and
comparing said correlation value with a threshold correlation value; and
wherein said classifying step includes the step of classifying said broadcast information as similar

to only the second stage reference pattern whose corresponding correlation value exceeds said threshold reference value.

13. A method according to claim 1 further including the training steps of:
(a) analyzing reference broadcast information to identify a spectrally distinctive portion thereof;
(b) determining a figure of merit for said distinctive portion using a peak value and a standard deviation value from said distinctive portion;
(c) generating a first stage reference pattern from said distinctive portion;
(d) testing the, generated first stage reference pattern for time alignment sensitivity;
(e) repeating said steps (a)-(d) when the generated first stage reference pattern does not pass the time sensitivity test; and
(f) using said spectrally distinctive portion to provide first and second stage reference patterns when said generated first stage reference pattern passes the time sensitivity test.

14. Apparatus for classifying broadcast information, comprising:
means (2,4) for receiving broadcast information;
processing means (18) for correlating the broadcast information with a library of second stage reference patterns (34); and
processing means (24;26) for classifying said broadcast information as similar to one of said second stage reference patterns based on said correlation,
characterized by
processing means (24;26;) for comparing said information with a library of first stage reference patterns (34); and
processing means (24;26) for queueing the first stage reference patterns in an order of their similarity to said information, said processing means (18) for correlating being supplied with said library of second stage reference patterns, which correspond to said first stage reference patterns in the queueing order established.

15. Apparatus according to claim 14 further including means for generating a plurality of analyzed waveforms from said broadcast information;
and wherein said processing means (24) includes processing means for (a1) time warping one of said analyzed waveforms to provide at least one time warped waveform; and wherein said processing means for (c) correlating includes processing means (18) for correlating both said one analyzed waveform and said time warped waveform with said library of second stage reference patterns.

16. Apparatus according to claim 15 wherein said processing means for (a1) time warping includes processing means for linearly time warping said one analyzed waveform to provide a stretched waveform.

17. Apparatus according to claim 14 wherein said means (4) for receiving includes means for simultaneously receiving a plurality of broadcast information, and wherein said processing means (24) includes means for performing the processing functions (a)-(d) on said plurality of broadcast information, substantially simultaneously.

18. Apparatus according to claim 14 wherein said processing means for (a) comparing includes processing means for (a1) processing said broadcast information to provide a plurality N of analyzed signal patterns corresponding to said broadcast information; (a2) generating a spectragram from said analyzed signal patterns; (a3) providing said spectragram to said processing means for (a) comparing for comparison to said first stage reference patterns; and (a4) providing one of said analyzed signal patterns to said processing means for (c) correlating for correlation with said second stage reference patterns.

19. Apparatus according to claim 18 wherein said processing means for (a) comparing includes processing means (6) for (a1a) bandpass filtering said broadcast information into a plurality of bands; and (a1b) computing a plurality of linear combinations of said plurality of bands to provide said plurality of analyzed signal patterns.

20. Apparatus according to claim 19 wherein said means (6) for (a1a) bandpass filtering includes means for

rectifying (10), and then lowpass filtering (12) each said band.

21. Apparatus according to claim 18 wherein said processing means for (a2) generating a spectragram includes processing means for (a2a) sampling said plurality of analyzed signal patterns at a predetermined rate; (a2b) constructing a time/frequency matrix having N frequency channels, n time channels, and a plurality of N(N) of matrix elements;

(a2c) calculating, for each matrix element, a matrix value $X_{N,n}$ as follows:

$$X_{N,n}(t) = [K_I \times V_{An}(t)] + \ldots [K_N \times V_{Nn}(t)]$$

where: t equals time at which samples are taken from said plurality of analyzed signal patterns;
V_{An} through V_{Nn} are amplitude values of first through N^{th} analyzed signal patterns taken at sample time t; and
K_I through K_N are constants preselected to minimize the influence of broadband impulsive energy.

22. Apparatus according to claim 21 wherein said processing means for (a2) generating a spectragram includes processing means for normalizing said matrix to provide said spectragram.

23. Apparatus according to claim 21 wherein each said first stage reference patterns comprises an Nxn reference matrix of elements, and wherein said processing means for (a) comprising includes processing means for (a5) measuring a variation between each element of the time/frequency matrix and a corresponding element of each of said reference matrices; (a6) summing the measured variations between the time/frequency matrix and each reference matrix to provide a distance measurement for each reference matrix; (a7) comparing each distance measurement with a threshold value; and (a8) eliminating those first stage reference patterns whose corresponding distance measurement exceeds said threshold value.

24. Apparatus according to claim 23 wherein said processing means for (b) queueing includes processing means for queueing non-eliminated first stage reference patterns according to their corresponding distance measurements.

25. Apparatus according to claim 14 wherein said processing means for (c) correlating includes processing means for (c1) calculating a correlation value for at least one second stage reference pattern with reference to said broadcast information; and (c2) comparing said correlation value with a threshold correlation value; and wherein said processing means for (d) classifying includes processing means for classifying said broadcast information as similar to only the second stage reference pattern whose corresponding correlation value exceeds said threshold reference value.

26. Apparatus according to claim 14 wherein said processing means includes processing means for (e) analyzing said broadcast information to identify a spectrally distinctive portion thereof; (f) determining a figure of merit for said distinctive portion using a peak value and a peak value standard deviation from said distinctive portion; (g) generating a first stage reference pattern from said distinctive portion; (h) testing the generated first stage reference pattern for time alignment sensitivity; (i) repeating the functions (e)-(h) when the generated first stage reference pattern does not pass the time sensitivity test; and (j) using said spectrally distinctive portion to provide first and second stage reference patterns when said generated first stage reference pattern passes the time sensitivity test.

Patentansprüche

1. Verfahren zur Klassifizierung von Rundfunkinformationen mit den Verfahrensschritten:
Empfang von Rundfunkinformation;
Korrelieren (S270) der Rundfunkinformation mit einer Bibliothek aus Bezugsmustern einer zweiten Stufen;
und auf der Basis dieses Korrelierens durchzuführendes Klassifizieren (S270,S300) dieser Rundfunkinformation hinsichtlich ihrer Ähnlichkeit mit einem der Bezugsmuster der zweiten Stufe,
gekennzeichnet dadurch,

daS zwischen Empfangen und Korrelieren diese Rundfunkinformation mit einer Bibliothek aus Bezugsmustern einer ersten Stufe verglichen wird (S150), wobei diese Bezugsmuster der ersten Stufe in einer Reihenfolge (S170) nach ihrer Ähnlichkeit mit dieser Rundfunkinformation eingeordnet werden; und daS diese Bibliothek der Bezugsmuster der zweiten Stufe den Bezugsmustern der ersten Stufe in der Reihenfolge entspricht, die in dem Verfahrensschritt der Einordnung in Reihenfolge erstellt worden ist.

2. Verfahren nach Anspruch 1 mit den weiteren Verfahrensschritten:
Erzeugen einer Vielzahl von analysierten Wellenformen aus dieser Rundfunkinformation;
Zeitverzerrung einer dieser analysierten Wellenformen, um wenigstens eine zeitverzerrte Wellenform zu erhalten;
wobei der Verfahrensschritt des Korrelierens das Korrelieren sowohl der einen analysierten Wellenform als auch dieser zeitverzerrten Wellenform mit der Bibliothek der Bezugsmuster der zweiten Stufe einschließt.

3. Verfahren nach Anspruch 2,
bei dem der Verfahrensschritt der Zeitverzerrung lineares Zeitverzerren dieser einen analysierten Wellenform einschließt, um eine gestreckte Wellenform zu erhalten.

4. Verfahren nach Anspruch 1,
wobei der Verfahrensschritt des Empfangens das gleichzeitige Empfangen einer Vielzahl von Rundfunkinformationen einschließt und wobei der Verfahrensschritt des Vergleichens, des Einordnens in einer Reihenfolge, des Korrelierens und des Klassifizierens hinsichtlich dieser Vielzahl von Rundfunkinformationen im wesentlichen gleichzeitig ausgeführt wird.

5. Verfahren nach Anspruch 1 mit den weiteren Verfahrensschritten:
Verarbeiten der Rundfunkinformation, um eine Vielzahl (N) analysierter Signalmuster zu erhalten, die dieser Rundfunkinformation entsprechen;
Erzeugen eines Spektrogramms dieser analysierten Signalmuster; Verwenden dieses Spektrogramms in dem Verfahrensschritt des Vergleichens, um einen Vergleich mit den Bezugsmustern der ersten Stufe durchzuführen und
Verwenden eines dieser analysierten Signalmuster im Verfahrensschritt des Korrelierens, um Korrelation mit den Bezugsmustern der zweiten Stufe durchzuführen.

6. Verfahren nach Anspruch 5,
wobei dieser Verfahrensschritt des Verarbeitens die folgenden Verfahrensschritte einschließt:
Bandpaßfiltern dieser Rundfunkinformation in eine Vielzahl von Bändern; und
Errechnen einer Vielzahl linearer Kombinationen dieser Vielzahl von Bändern, um die Vielzahl analysierter Signalmuster vorzusehen.

7. Verfahren nach Anspruch 6,
wobei dieser Verfahrensschritt des Bandpaßfilterns das Bandpaßfiltern, Gleichrichten und dann Tiefpaßfiltern eines jeden Bandes einschließt.

8. Verfahren nach Anspruch 5,
wobei der Verfahrensschritt des Erzeugens eines Spektrogramms die folgenden Verfahrensschritte einschließt:
Abtasten einer Vielzahl analysierter Signalmuster mit einer vorgegebenen Abtastrate;
Aufstellen einer Zeit/Frequenz-Matrix mit N Frequenzkanälen, mit n Zeitkanälen und mit einer Vielzahl N(n) Matrixelementen;
Errechnen aus einem jeden Matrixelement einen Matrixwert $X_{N,n}$ wie folgt:

$$X_{N,n}(t) = (K_I \times V_{An}(t)) + \ldots (K_N \times V_{Nn}(t))$$

worin:
t der Zeitpunkt ist, zu dem die Abtastproben dieser Vielzahl analysierter Signalmuster genommen worden sind;
V_{An} bis V_{Nn} die Amplitudenwerte der ersten bis N-ten analysierten Signalmuster sind, die zur Abtastzeit t abgetastet sind; und

K_1 bis K_N voraus ausgewählte Konstanten sind, mit denen der Einfluß von Breitband-Impuls-Störenergie minimiert wird.

9. Verfahren nach Anspruch 8,
wobei weiterhin ein Normieren dieser Matrix eingeschlossen ist, um das Spektrogramm zu schaffen.

10. Verfahren nach Anspruch 8,
wobei ein jedes der Bezugsmuster der ersten Stufe eine N x n-Bezugsmatrix aus Elementen umfaßt und wobei das Vergleichen die folgenden Verfahrensschritte umfaßt:
Messen einer Variation zwischen einem jeden Element der Zeit/Frequenz-Matrix und einem korrespondierenden Element einer jeden dieser Bezugsmatrizen;
Aufsummieren der gemessenen Variationen zwischen der Zeit/Frequenz-Matrix und einer jeden Bezugsmatrix, um einen Abstandsmeßwert für eine jede Bezugsmatrix zu schaffen;
Vergleichen eines jeden Abstandsmeßwertes mit einem Schwellenwert; und
Ausschließen solcher Bezugsmuster der ersten Stufe, deren entsprechender Abstandsmeßwert diesen Schwellenwert übersteigt.

11. Verfahren nach Anspruch 10,
wobei der Verfahrensschritt des Einordnens in die Reihenfolge den Verfahrensschritt des Einordnens nicht ausgeschlossener Bezugsmuster der ersten Stufe entsprechend ihrer Abstandsmeßwerte beinhaltet.

12. Verfahren nach Anspruch 1,
wobei das Korrelieren die folgenden Verfahrensschritte einschließt: Errechnen eines Korrelationswertes für wenigstens ein Bezugsmuster der zweiten Stufe in Bezug auf die Rundfunkinformation; und
Vergleichen dieses Korrelationswertes mit einem Korrelations-Schwellenwert;
wobei dieser Verfahrensschritt des Klassifizierens das Klassifizieren dieser Rundfunkinformation hinsichtlich Ähnlichkeit mit nur demjenigen Bezugsmuster der zweiten Stufe einschließt, dessen entsprechender Korrelationswert diesen Bezugs-Schwellenwert übersteigt.

13. Verfahren nach Anspruch 1,
wobei des weiteren die folgenden Übungsschritte eingeschlossen sind:
a) Analysieren einer Referenz-Rundfunkinformation, um einen spektral klar erkennbaren Anteil derselben zu identifizieren;
b) Bestimmen einer Güteziffer für diesen klar erkennbaren Anteil unter Benutzung eines Scheitelwertes und eines Standard-Abweichungswertes aus diesem klar erkennbaren Anteil;
c) Erzeugung eines Bezugsmusters der ersten Stufe aus diesem klar erkennbaren Anteil;
d) Testen des Bezugsmusters der ersten Stufe für Zeitabgleichempfindlichkeit;
e) Wiederholen dieser Verfahrensschritte a) bis d) dann, wenn das erzeugte Bezugsmuster der ersten Stufe den Zeitempfindlichkeitstest nicht passiert; und
f) Benutzen des spektral klar erkennbaren Anteils, um Bezugsmuster der ersten und der zweiten Stufe dann vorzusehen, wenn erzeugte Bezugsmuster der ersten Stufe den Zeitempfindlichkeitstest passieren.

14. Vorrichtung zum Klassifizieren von Rundfunkinformation mit:
Mitteln (2,4) zum Empfangen von Rundfunkinformation;
Prozessormittel (18) zum Korrelieren der Rundfunkinformation mit einer Bibliothek aus Bezugsmustern (34) der zweiten Stufe; und
Prozessormittel (24,26) zum Klassifizieren dieser Rundfunkinformation hinsichtlich Ähnlichkeit mit einem dieser Bezugsmuster der zweiten Stufe, beruhend auf dieser Korrelation;
gekennzeichnet dadurch,
Prozessormittel (24,26) zum Vergleichen dieser Information mit einer Bibliothek aus Bezugsmustern (34) einer ersten Stufe und Prozessormittel (24;26) zum Einordnen der Bezugsmuster der ersten Stufe in eine Reihenfolge hinsichtlich deren Ähnlichkeit mit dieser Information, wobei die Prozessormittel (18) für das Korrelieren mit der Bibliothek der Bezugsmuster der zweiten Stufe gespeist werden, die den Bezugsmustern der ersten Stufe, diese in der Reihenfolge aufgestellt, entsprechen.

15. Vorrichtung nach Anspruch 14,

aie des weiteren Mittel zur Erzeugung einer Vielzahl analysierter Wellenformen der Rundfunkinformation einschließt;
und wobei diese Prozessormittel (24) Prozessormittel (a1) zur Zeitverzerrung einer dieser analysierten Wellenformen einschließen, um wenigstens eine zeitverzerrte Wellenform zu erzeugen, und wobei diese Prozessormittel zur (c) Korrelation Prozessormittel (18) zum Korrelieren sowohl dieser einen analysierten Wellenform als auch der zeitverzerrten Wellenform mit dieser Bibliothek aus Bezugsmustern der zweiten Stufe einschließen.

16. Vorrichtung nach Anspruch 15,
bei der diese Prozessormittel zum (a1) Zeitverzerren Prozessormittel für lineares Zeitverzerren dieser einen analysierten Wellenform einschließen, um eine gestreckte Wellenform zu erzeugen.

17. Vorrichtung nach Anspruch 14,
bei der diese Mittel (4) zum Empfang Mittel zum gleichzeitigen Empfangen einer Vielzahl von Rundfunkinformationen einschließen und bei der diese Prozessormittel (24) Mittel zum Ausführen der Prozessorfunktionen/Schritte (a) - (d) einschließen, die bezüglich der Vielzahl der Rundfunkinformationen im wesentlichen gleichzeitig ausgeführt werden.

18. Vorrichtung nach Anspruch 14,
bei der diese Prozessormittel zum (a) Vergleichen Prozessormittel einschließen zum
(a1) Verarbeiten dieser Rundfunkinformation, um eine Vielzahl N analysierter Signalmuster zu erzeugen, die dieser Rundfunkinformation entsprechen;
(a2) Erzeugen eines Spektrogramms aus diesen analysierten Signalmustern;
(a3) Zuführen dieses Spektrogramms diesen Prozessormitteln zum (a) Vergleichen, um Vergleich mit den Bezugsmustern der ersten Stufe durchzuführen und
(a4) Zuführen eines der analysierten Signalmuster den Prozessormitteln zum (c) Korrelieren, um die Korrelation mit den Bezugsmustern der zweiten Stufe ourchzuführen.

19. Vorrichtung nach Anspruch 18,
bei der diese Prozessormittel zum (a) Vergleichen Prozessormittel (6) einschließen zum
(a1a) Bandpaßfiltern dieser Rundfunkinformation in eine Vielzahl Bänder; und
(a1b) Errechnen einer Vielzahl linearer Kombinationen dieser Vielzahl Bänder, um die Vielzahl der analysierten Signalmuster zu erhalten.

20. Vorrichtung nach Anspruch 19,
bei der diese Mittel (6) zum (a1a) Bandpaßfiltern Mittel zum Gleichrichten (10) und dann Tiefpaßfiltern (12) eines jeden Bandes einschließen.

21. Vorrichtung nach Anspruch 18,
bei der diese Prozessormittel zum (a2) Erzeugen eines Spektrogramms Prozessormittel einschließen zum
(a2a) Abtasten dieser Vielzahl analysierter Signalmuster mit einer vorgegebenen Abtastrate;
(a2b) Aufstellen einer Zeit/Frequenz-Matrix mit N Frequenzkanälen, n Zeitkanälen und einer Vielzahl N-(n) Matrixelementen; und
(a2c) Errechnen eines Matrixwertes ($X_{N,n}$) für ein jedes Matrixelement in der folgenden Weise:

$$X_{N,n}(t) = (K_I \times V_{An}(t)) + \ldots (K_N \times V_{Nn}(t))$$

worin:

t	der Zeitpunkt ist, zu dem die Abtastproben dieser Vielzahl analysierter Signalmuster genommen worden sind;
V_{An} bis V_{Nn}	die Amplitudenwerte der ersten bis N-ten analysierten Signalmuster sind, die zur Abtastzeit t abgetastet sind; und
K_I bis K_N	voraus ausgewählte Konstanten sind, mit denen der Einfluß von Breitband-Impuls/Störenergie minimiert wird.

22. Vorrichtung nach Anspruch 21,
bei der die Prozessormittel zum (a2) Erzeugen eines Spektromgramms Prozessormittel zum Normieren

dieser Matrix einschließen, um das Spektrogramm zu bilden.

23. Vorrichtung nach Anspruch 21,
bei der ein jedes der Bezugsmuster der ersten Stufe eine N x n-Bezugsmatrix aus Elementen umfaßt und wobei die Prozessormittel zum
(a) Vergleichen Prozessormittel einschließen zum
(a5) Messen einer Variation zwischen einem jeden Element der Zeit/Frequenz-Matrix und einem korrespondierenden Element einer jeden dieser Bezugsmatrizen;
(a6) Aufsummieren der gemessenen Variationen zwischen der Zeit/Frequenz-Matrix und einer jeden Bezugsmatrix, um einen Abstandsmeßwert für eine jede Bezugsmatrix zu schaffen;
(a7) Vergleichen eines jeden Abstandsmeßwertes mit einem Schwellenwert; und
(a8) Ausschließen solcher Bezugsmuster der ersten Stufe, deren entsprechender Abstandsmeßwert diesen Schwellenwert übersteigt.

24. Vorrichtung nach Anspruch 23,
bei der die Prozessormittel zum (b) Einordnen in die Reihenfolge Prozessormittel zum Einordnen nicht ausgeschlossener Bezugsmuster der ersten Stufe entsprechend ihrer Abstandswerte einschließt.

25. Vorrichtung nach Anspruch 14,
bei der die Prozessormittel zum (c) Korrelieren Prozessormittel einschließen zum
(c1) Errechnen eines Korrelationswertes für wenigstens ein Bezugsmuster der zweiten Stufe in Bezug auf die Rundfunkinformation; und
(c2) Vergleichen dieses Korrelationswertes mit einem Korrelations-Schwellenwert und
wobei dieser Verfahrensschritt (d) zum Klassifizieren das Klassifizieren dieser Rundfunkinformation hinsichtlich Ähnlichkeit mit nur demjenigen Bezugsmuster der zweiten Stufe einschließt, dessen entsprechender Korrelationswert diesen Bezugs-Schwellenwert übersteigt.

26. Vorrichtung nach Anspruch 14,
bei der diese Prozessormittel Prozessormittel einschließen zum
(e) Analysieren dieser Rundfunkinformation, um einen spektral klar erkennbaren Anteil derselben zu identifizieren;
(f) Bestimmen einer Güteziffer für diesen klar erkennbaren Anteil unter Benutzung eines Scheitelwertes und einer Scheitelwert-Standardabweichung aus diesem klar erkennbaren Anteil;
(g) Erzeugen eines Bezugsmusters der ersten Stufe aus diesem klar erkennbaren Anteil;
(h) Testen des Bezugsmusters der ersten Stufe für Zeitabgleichempfindlichkeit;
(i) Wiederholen dieser Funktionen (e) - (h) dann, wenn das erzeugte Bezugsmuster der ersten Stufe den Zeitempfindlichkeitstest nicht passiert; und
(j) Benutzen des spektral klar erkennbaren Anteils, um Bezugsmuster der ersten und der zweiten Stufe dann vorzusehen, wenn erzeugte Bezugsmuster der ersten Stufe den Zeitempfindlichkeitstest passieren.

Revendications

1. Procédé de classification d'information diffusée, comprenant les étapes de :
réception d'une information diffusée
corrélation (S270) de l'information diffusée à une bibliothèque de motifs de référence de second niveau ; et
classification (S270, S300) de ladite information diffusée comme étant similaire à l'un desdits motifs de référence de second niveau sur la base de ladite étape de corrélation ;
caractérisé en ce que :
entre la réception et la corrélation, ladite information diffusée est comparée (S150) à une bibliothèque de motifs de référence de premier niveau, lesdits motifs de référence de premier niveau étant mis en file (S170) selon un ordre correspondant à leur similarité par rapport à ladite information diffusée : et
en ce que ladite bibliothèque de motifs de référence de second niveau correspond auxdits motifs de référence de premier niveau suivant l'ordre de file établi lors de l'étape de mise en file.

2. Procédé selon la revendication 1, incluant en outre les étapes de :

génération d'une pluralité de formes d'onde analysées à partir de ladite information diffusée ;

modification temporelle d'une desdites formes d'onde analysées pour produire au moins une forme d'onde modifiée temporellement ; et

dans lequel ladite étape de corrélation inclut l'étape de corrélation à la fois de ladite une forme d'onde analysèe et de ladite forme d'onde modifiée temporellement à ladite bibliothèque de motifs de référence de second niveau.

3. Procédé selon la revendication 2, dans lequel ladite étape de modification temporelle inclut l'étape de modification temporelle linéaire de ladite forme d'onde analysée pour produire une forme d'onde étirée.

4. Procédé selon la revendication 1 dans lequel ladite étape de réception inclut l'étape de réception simultanée d'une pluralité d'informations diffusées et dans lequel lesdites étapes de comparaison, de mise en file, de corrélation et de classification sont effectuées sur ladite pluralité d'informations diffusées, sensiblement simultanément.

5. Procédé selon la revendication 1, incluant en outre les étapes de :

traitement de ladite information diffusée pour produire une pluralité N de motifs de signal analysés correspondant à ladite information diffusée ;

génération d'un spectrogramme à partir desdits motifs de signal analysés ;

production dudit spectrogramme pour ladite étape de comparaison afin d'effectuer une comparaison par rapport auxdits motifs de référence de premier niveau ; et

production de l'un desdits motifs de signal analysés pour ladite étape de corrélation afin d'effectuer une corrélation avec lesdits motifs de référence de second niveau.

6. Procédé selon la revendication 5 dans lequel ladite étape de traitement inclut les étapes de :

filtrage passe-bande de ladite information diffusèe selon une pluralité de bandes ; et

calcul d'une pluralité de combinaisons linéaires de ladite pluralité de bandes pour produire ladite pluralité de motifs de signal analysés.

7. Procédé selon la revendication 6, dans lequel ladite étape de filtrage passe-bande inclut les étapes de filtrage passe-bande, le redressage puis le filtrage passe-bas de chaque dite bande.

8. Procédé selon la revendication 5, dans lequel ladite étape de génération d'un spectrogramme inclut les étapes de :

échantillonnage de ladite pluralité de motifs de signal analysés selon un débit prédéterminé

construction d'une matrice temps/fréquence comportant N canaux de fréquence, n canaux temporels et une pluralité N(n) d'éléments matriciels ;

calcul, pour chaque élément matriciel, d'une valeur matricielle $X_{N,n}$ comme suit :

$$X_{N,n}(t) = [K_1 \times V_{An}(t)] + \ldots [K_N \times V_{Nn}(t)]$$

où : t est un instant auquel des échantillons sont pris parmi ladite pluralité de motifs de signal analysés

V_{An} à V_{Nn} sont des valeurs d'amplitude des premier à N-ième motifs de signal analysés prises à un instant d'échantillonnage t : et

K_1 à K_N sont des constantes présélectionnées afin de minimiser l'influence d'une énergie impulsive de diffusion.

9. Procédé selon la revendication 8, incluant en outre l'étape de normalisation de ladite matrice pour produire ledit spectrogramme.

10. Procédé selon la revendication 8, dans lequel chaque dit motif de référence de premier niveau comprend une matrice d'éléments de référence N x n et dans lequel ladite étape de comparaison comprend les étapes de :

mesure d'une variation entre chaque élément de la matrice temps/fréquence et un élément correspondant de chacune desdites matrices de référence ;

sommation des variations mesurées entre la matrice temps/fréquence et chaque matrice de référence pour produire une mesure de distance pour chaque matrice de référence ;

comparaison de chaque mesure de distance à une valeur de seuil ; et

élimination des motifs de référence de premier niveau dont la mesure de distance correspondante excède ladite valeur de seuil.

11. Procédé selon la revendication 10, dans lequel ladite étape de mise en file comprend l'étape de mise en ordre de motifs de référence de premier niveau non éliminés conformément à leurs mesures de distance correspondantes.

12. Procédé selon la revendication 1, dans lequel ladite étape de corrélation inclut les étapes de :
calcul d'une valeur de corrélation pour au moins un motif de référence de second niveau par référence à ladite information diffusée ; et
comparaison de ladite valeur de corrélation à une valeur de corrélation de seuil ; et
dans lequel ladite étape de classification inclut l'étape de classification de ladite information diffusée comme étant similaire à seulement le motif de référence de second niveau dont la valeur de corrélation correspondante excède ladite valeur de référence de seuil.

13. Procédé selon la revendication 1, incluant en outre les étapes d'apprentissage suivantes :
(a) analyse d'une information diffusée de référence afin d'identifier une partie spectralement distinctive de celle-ci ;
(b) détermination d'une figure de mérite pour ladite partie distinctive an utilisant une valeur de pic et une valeur de déviation standard à partir de ladite partie distinctive ;
(c) génération d'un motif de référence de premier niveau à partir de ladite partie distinctive ;
(d) test du motif de référence de premier niveau généré du point de vue de la sensibilité d'alignement temporel ;
(e) répétition desdites étapes (a) à (d) lorsque le motif de référence de premier niveau généré ne satisfait pas au test de sensibilité temporelle ; et
(f) utilisation de ladite partie spectralement distinctive pour produire des motifs de référence de premier et de second niveaux lorsque ledit motif de référence de premier niveau généré satisfait au test de sensibilité temporelle.

14. Appareil de classification d'information diffusée comprenant :
un moyen (2, 4) pour recevoir une information diffusée
un moyen de traitement (18) pour corréler l'information diffusée à une bibliothèque de motifs de référence de second niveau (34); et
un moyen de traitement (24 ; 26) pour classifier ladite information diffusée comme étant similaire à l'un desdits motifs de référence de second niveau sur la base de ladite corrélation ;
caractérisé par :
un moyen de traitement (24 ; 26) pour comparer ladite information à une bibliothèque de motifs de référence de premier niveau (34) ; et
un moyen de traitement (24 ; 26) pour mettre en file les motifs de référence de premier niveau selon l'ordre de leur similarité par rapport à ladite information, ledit moyen de traitement (18) de corrélation se voyant appliquer ladite bibliothèque de motifs de référence de second niveau, lesquels correspondent auxdits motifs de référence de premier niveau selon l'ordre de file établi.

15. Appareil selon la revendication 14, incluant en outre un moyen pour générer une pluralité de formes d'onde analysées à partir de ladite information diffusée ;
et dans lequel ledit moyen de traitement (24) inclut un moyen de traitement de (a1) modification temporelle de l'une desdites formes d'onde analysées afin de produire au moins une forme d'onde modifiée temporellement et dans lequel ledit moyen de traitement de (c) corrélation inclut un moyen de traitement (18) pour corréler à la fois ladite une forme d'onde analysée et ladite forme d'onde modifiée temporellement à ladite bibliothèque de motifs de référence de second niveau.

16. Appareil selon la revendication 15, dans lequel ledit moyen de traitement de (a1) modification temporelle inclut un moyen de traitement pour modifier temporellement linéairement ladite une forme d'onde analysée afin de produire une forme d'onde étirée.

17. Appareil selon la revendication 14, dans lequel ledit moyen (4) de reception inclut un moyen pour recevoir simultanément une pluralité d'informations diffusées et dans lequel ledit moyen de traitement (24) inclut un moyen pour effectuer les fonctions de traitement (a) à (d) sur ladite pluralité d'informa-

tions diffusées, sensiblement simultanément.

18. Appareil selon la revendication 14, dans lequel ledit moyen de traitement pour (a) effectuer une comparaison inclut un moyen de traitement pour (a1) traiter ladite information diffusée pour produire une pluralité N de motifs de signal analysés correspondant à ladite information diffusée : pour (a2) générer un spectrogramme à partir desdits motifs de signal analysés ; pour (a3) produire ledit spectrogramme pour ledit moyen de traitement pour (a) effectuer une comparaison par rapport auxdits motifs de référence de premier niveau ; et pour (a4) produire l'un desdits motifs de signal analysés pour ledit moyen de traitement pour (c) effectuer une corrélation avec lesdits motifs de référence de second niveau.

19. Appareil selon la revendication 18, dans lequel ledit moyen de traitement pour (a) effectuer une comparaison inclut un moyen de traitement (6) pour (a1a) filtrer passe-bande ladite information diffusée selon une pluralité de bandes ; et pour (a1b) calculer une pluralité de combinaisons linéaires de ladite combinaison de bandes pour produire ladite pluralité de motifs de signal analysés.

20. Appareil selon la revendication 19, dans lequel ledit moyen (6) pour (a1a) filtrer passe-bande inclut un moyen pour redresser (10) puis pour filtrer passe-bas (12) chaque dite bande.

21. Appareil selon la revendication 18, dans lequel ledit moyen de traitement pour (a2) générer un spectrogramme inclut un moyen de traitement pour (a2a) échantillonner ladite pluralité de motifs de signal analysés selon un débit prédéterminé ; pour (a2b) construire une matrice temps/fréquence comportant N canaux de fréquence, n canaux temporels et une pluralité N(n) d'éléments matriciels ; pour (ac2) calculer pour chaque élément matriciel, une valeur matricielle $X_{N,n}$ comme suit :

$$X_{N,n}(t) = [K_1 \times V_{An}(t)] + \ldots [K_N \times V_{Nn}(t)]$$

où : t est un instant auquel des échantillons sont pris parmi ladite pluralité de motifs de signal analysés ;

V_{An} à V_{Nn} sont des valeurs d'amplitude des premier à N-ième motifs de signal analysés prises à un instant d'échantillonnage t ; et

K_1 à K_N sont des constantes présélectionnées afin de minimiser l'influence d'une énergie impulsive de diffusion.

22. Appareil selon la revendication 21, dans lequel ledit moyen de traitement pour (a2) générer un spectrogramme inclut un moyen de traitement pour normaliser ladite matrice pour produire ledit spectrogramme.

23. Appareil selon la revendication 21, dans lequel chacun desdits motifs de référence de premier niveau comprend une matrice d'éléments de référence N x n et dans lequel ledit moyen de traitement pour (a) effectuer une comparaison inclut un moyen de traitement pour (a5) mesurer une variation entre chaque élément de la matrice temps/fréquence et un élément correspondant de chacune desdites matrices de référence ; pour (a6) sommer les variations mesurées entre la matrice temps/fréquence et chaque matrice de référence pour produire une mesure de distance pour chaque matrice de référence ; pour (a7) comparer chaque mesure de distance à une valeur de seuil ; et pour (a8) éliminer les motifs de référence de premier niveau dont la mesure de distance correspondante excède ladite valeur de seuil.

24. Appareil selon la revendication 23, dans lequel ledit moyen de traitement pour (b) effectuer une mise en file inclut un moyen de traitement pour effectuer une mise en file des motifs de référence de premier niveau non éliminés conformément à leurs mesures de distance correspondantes.

25. Appareil selon la revendication 14, dans lequel ledit moyen de traitement pour (c) effectuer une corrélation inclut un moyen de traitement pour (c1) calculer une valeur de corrélation pour au moins un motif de référence de second niveau par référence à ladite information diffusée ; et pour (c2) comparer ladite valeur de corrélation à une valeur de corrélation de seuil ; et dans lequel ledit moyen de traitement pour (d) effectuer une classification inclut un moyen de traitement pour classifier ladite information diffusée comme étant similaire à seulement le motif de référence de second niveau dont la valeur de corrélation correspondante excède ladite valeur de référence de seuil.

26. Appareil selon la revendication 14, dans lequel ledit moyen de traitement inclut un moyen de traitement pour (e) analyser ladite information diffusée afin d'identifier une partie spectralement distinctive de celle-ci ; pour (f) déterminer une figure de mérite pour ladite partie distinctive en utilisant une valeur de pic et une déviation standard de valeur de pic à partir de ladite partie distinctive ; pour (g) générer un motif de référence de premier niveau à partir de ladite partie distinctive ; pour (h) tester le motif de référence de premier niveau généré du point de vue de la sensibilité d'alignement temporel ; pour (i) répéter les fonctions (e) à (h) lorsque le motif de référence de premier niveau généré ne satisfait pas au test de sensibilité temporelle ; et pour (j) utiliser ladite partie spectralement distinctive pour produire des motifs de référence de premier et second niveaux lorsque ledit motif de référence de premier niveau généré satisfait au test de sensibilité.

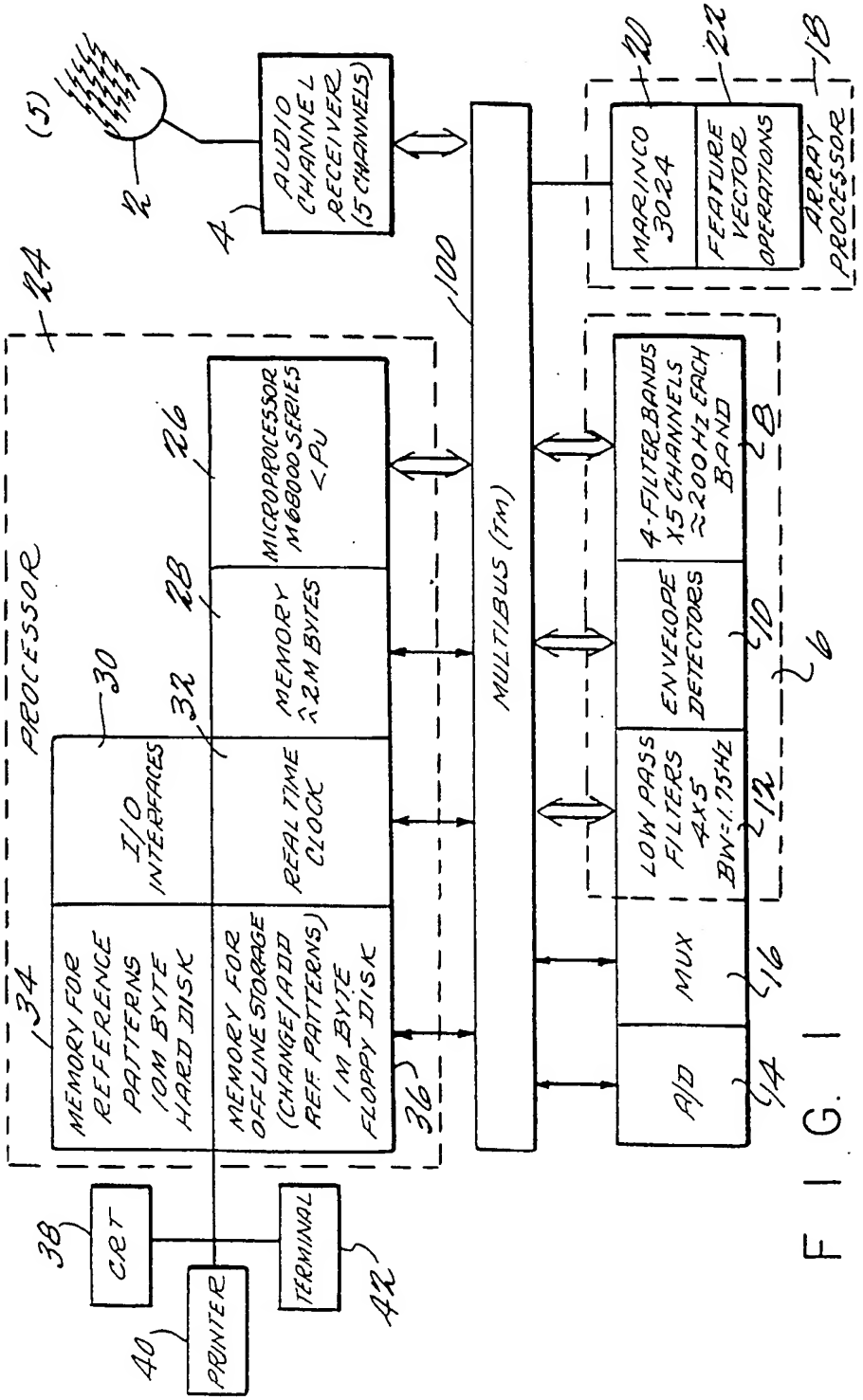

(5)
2
AUDIO CHANNEL RECEIVER (5 CHANNELS)
4
MARINCO 3024
FEATURE VECTOR OPERATIONS
ARRAY PROCESSOR
20
22
18
100
MULTIBUS (TM)
PROCESSOR
24
MICROPROCESSOR M68000 SERIES CPU
26
MEMORY ≈2M BYTES
28
REAL TIME CLOCK
32
I/O INTERFACES
30
MEMORY FOR REFERENCE PATTERNS (ON BYTE HARD DISK
34
MEMORY FOR OFFLINE STORAGE (CHANGE/ADD REF. PATTERNS) 1M BYTE FLOPPY DISK
36
4-FILTER BANDS X5 CHANNELS ≈200Hz EACH BAND
8
ENVELOPE DETECTORS
10
6
LOW PASS FILTERS 4X5 BW=1.75Hz
12
MUX
16
A/D
14
CRT
38
PRINTER
40
TERMINAL
42

FIG. 1

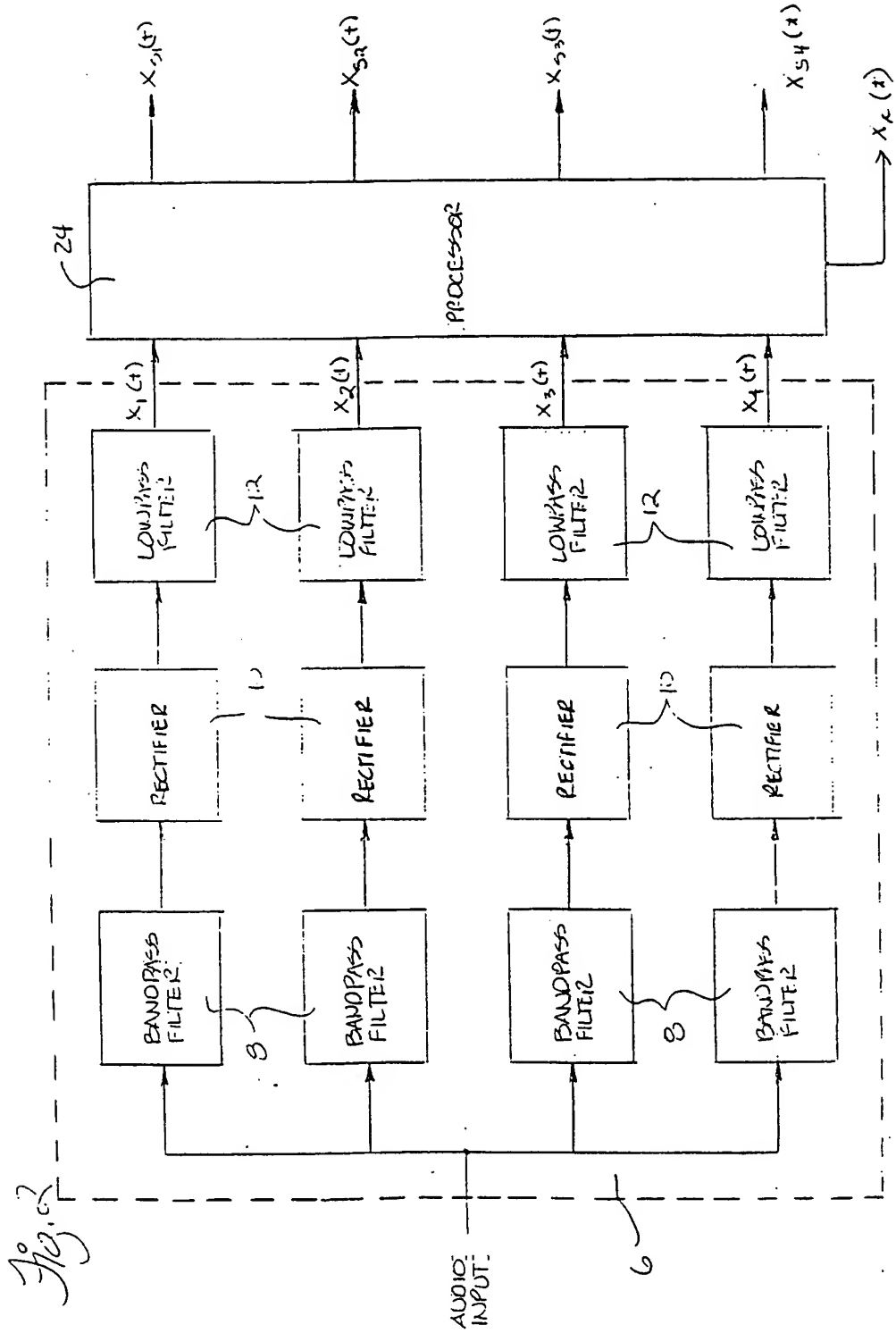

PROCESSOR
24
BANDPASS FILTER
RECTIFIER
LOWPASS FILTER
AUDIO INPUT
6
8
10
12

Fig. 2

FIG. 4

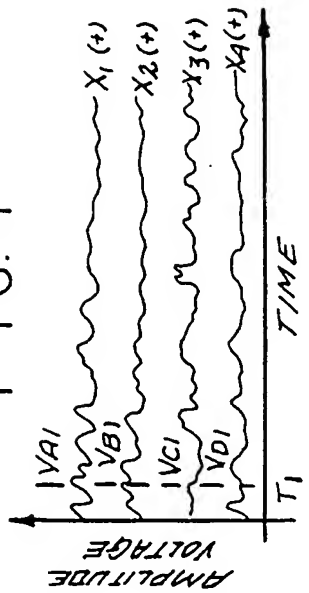


FIG. 5

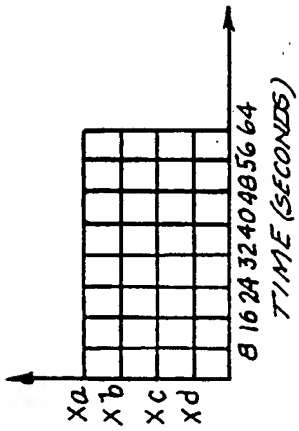
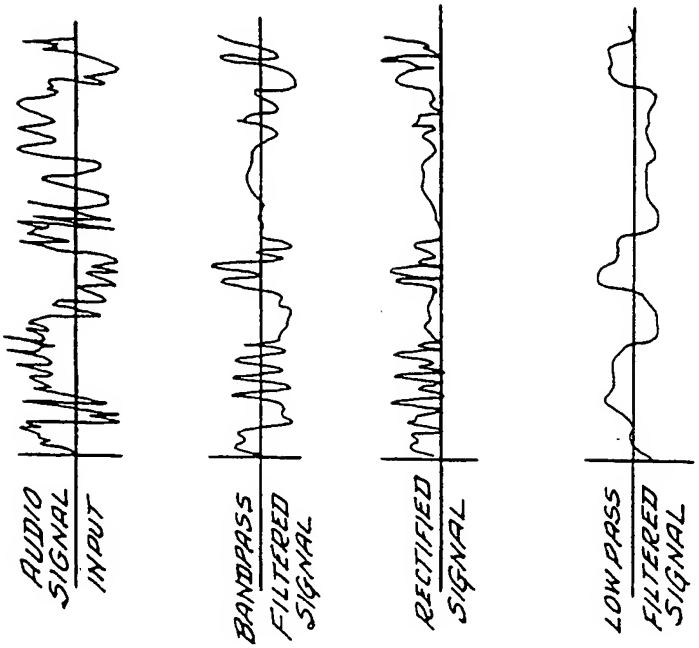


FIG. 3



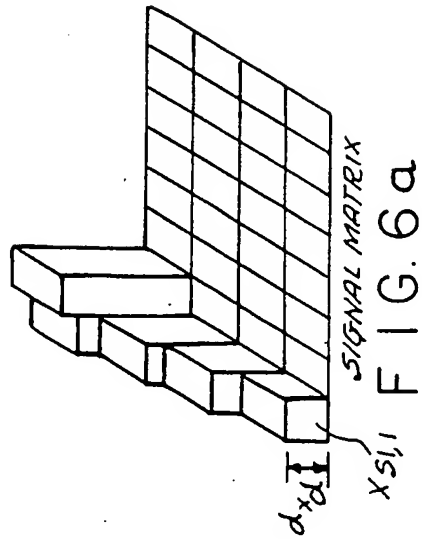


FIG. 6a

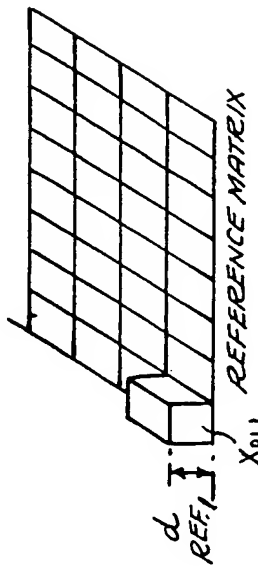


FIG. 6b

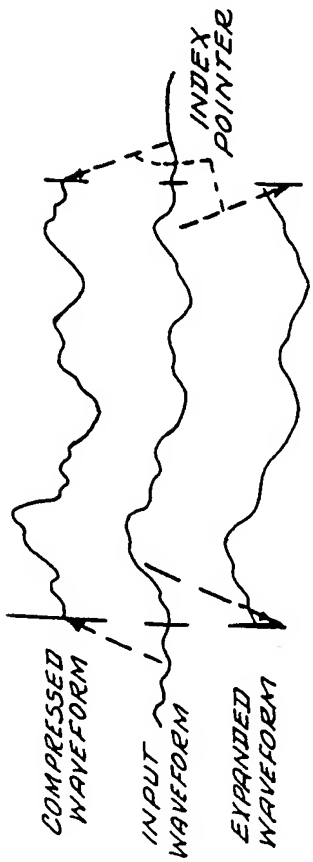


FIG. 7

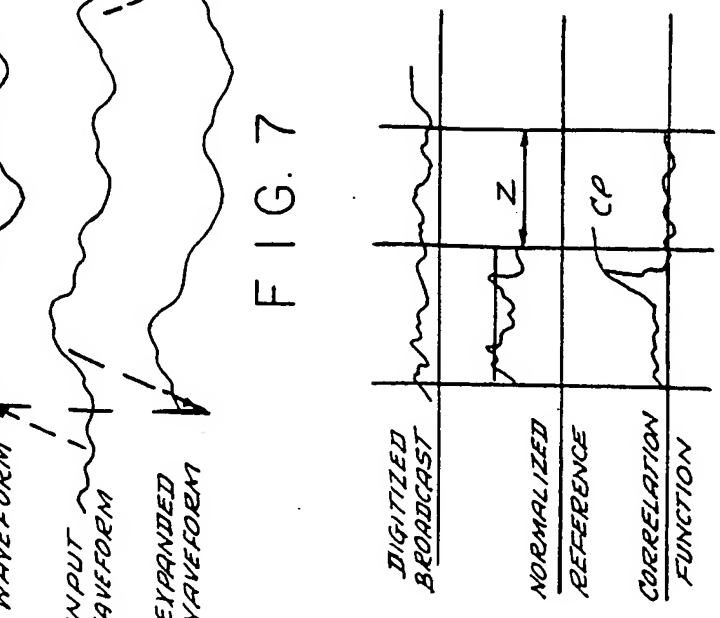


FIG. 8

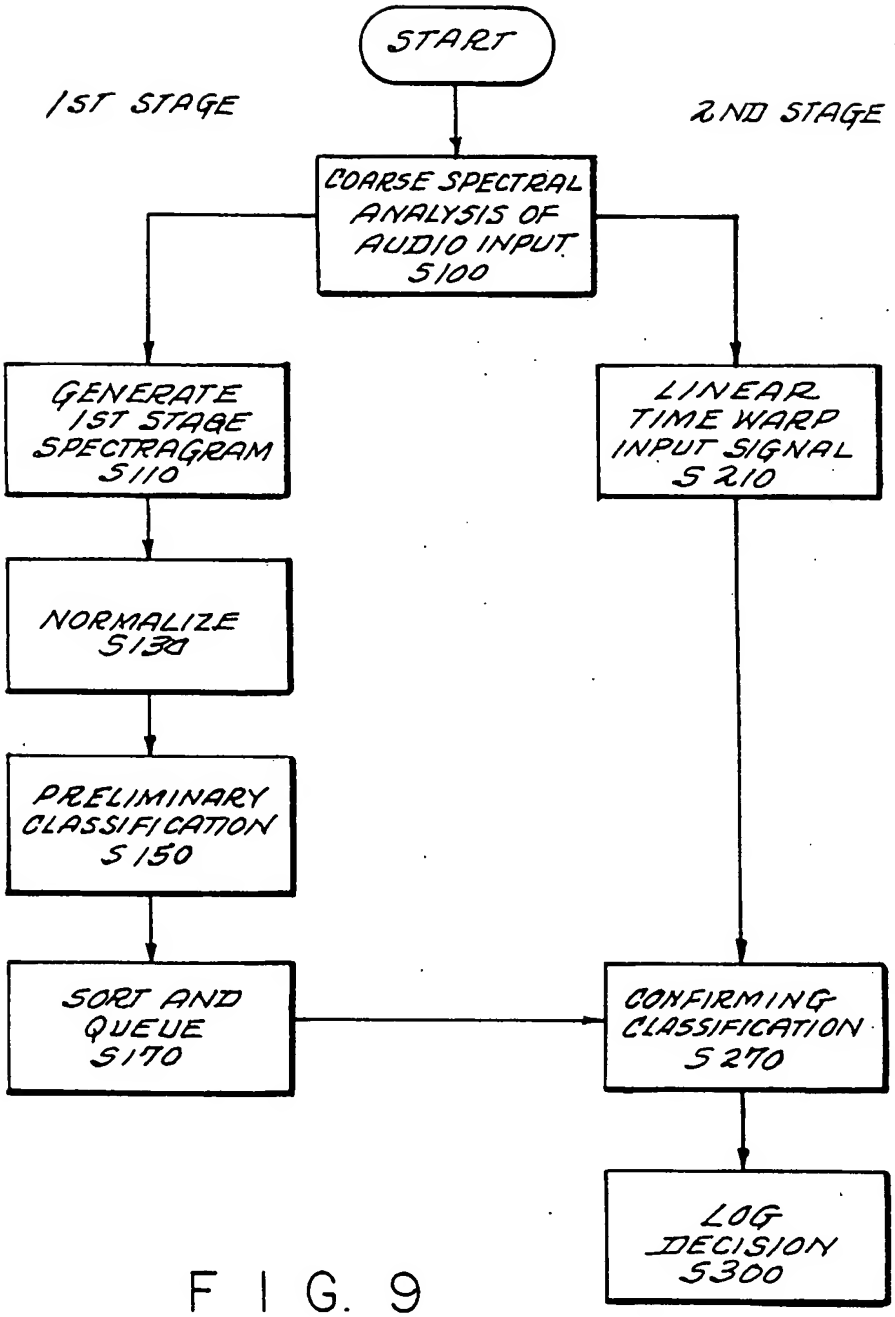

START
1ST STAGE
2ND STAGE
COARSE SPECTRAL ANALYSIS OF AUDIO INPUT S100
GENERATE 1ST STAGE SPECTRAGRAM S110
NORMALIZE S130
PRELIMINARY CLASSIFICATION S150
SORT AND QUEUE S170
LINEAR TIME WARP INPUT SIGNAL S210
CONFIRMING CLASSIFICATION S270
LOG DECISION S300

FIG. 9

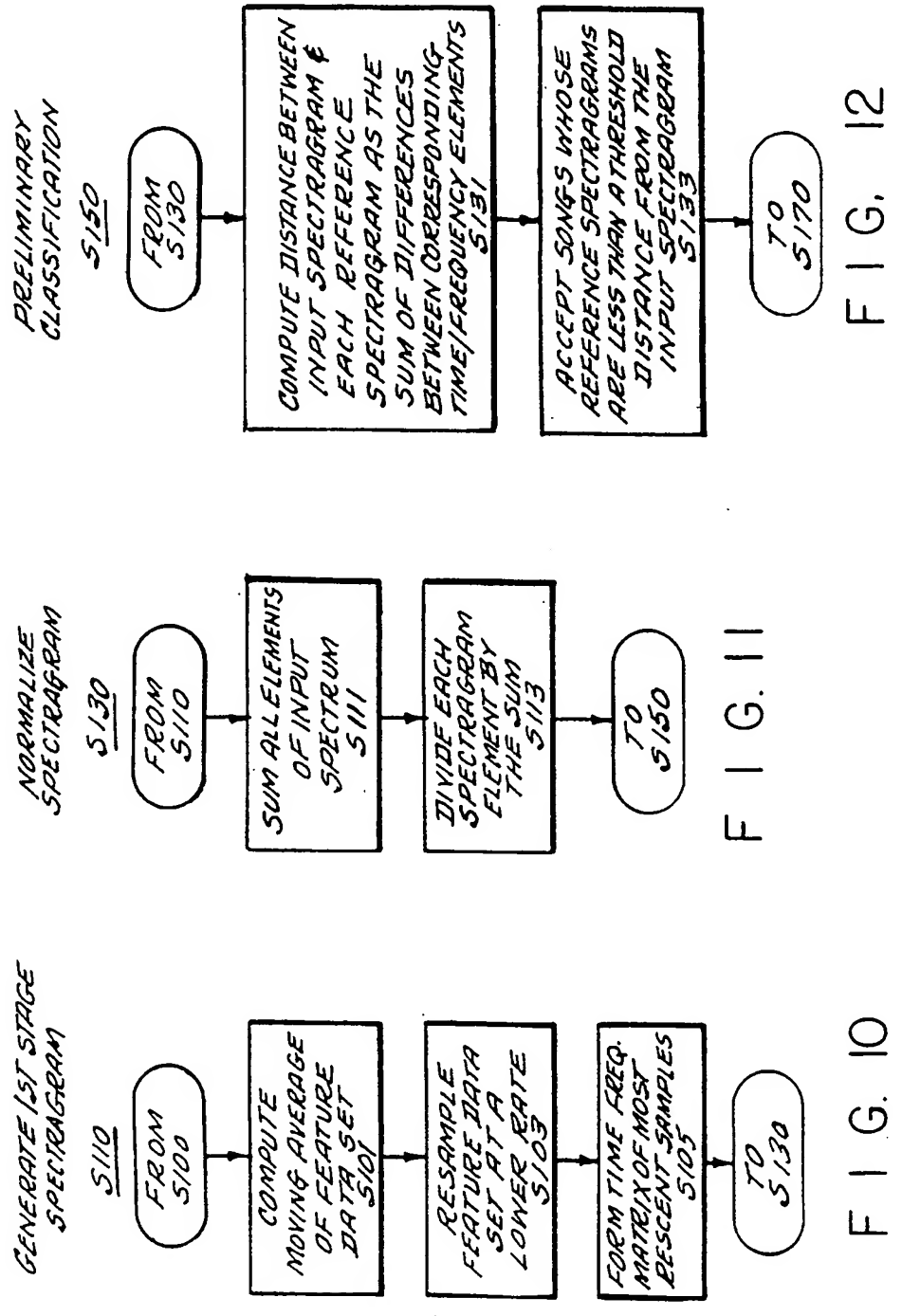
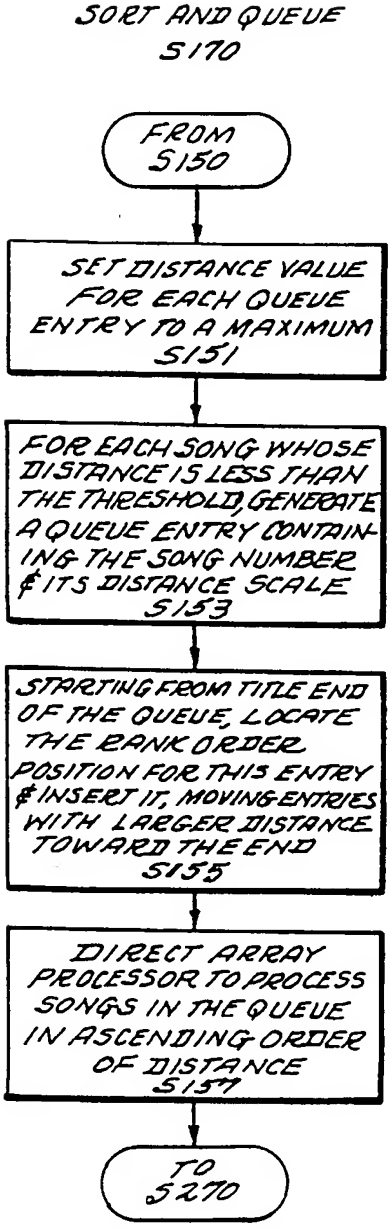


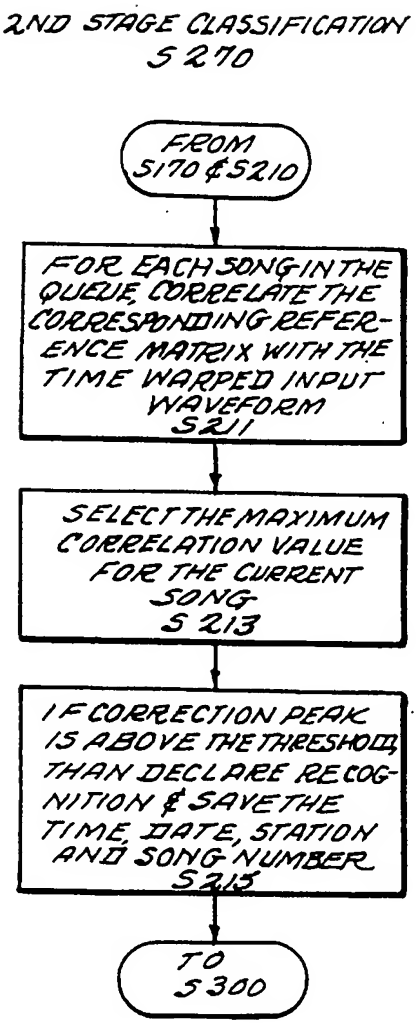
FIG. 10

FIG. 11

FIG. 12


SORT AND QUEUE
S170
FROM S150
SET DISTANCE VALUE FOR EACH QUEUE ENTRY TO A MAXIMUM S151
FOR EACH SONG WHOSE DISTANCE IS LESS THAN THE THRESHOLD, GENERATE A QUEUE ENTRY CONTAINING THE SONG NUMBER & ITS DISTANCE SCALE S153
STARTING FROM TITLE END OF THE QUEUE, LOCATE THE RANK ORDER POSITION FOR THIS ENTRY & INSERT IT, MOVING ENTRIES WITH LARGER DISTANCE TOWARD THE END S155
DIRECT ARRAY PROCESSOR TO PROCESS SONGS IN THE QUEUE IN ASCENDING ORDER OF DISTANCE S157
TO S270

F I G. 13


2ND STAGE CLASSIFICATION
S270
FROM S170 & S210
FOR EACH SONG IN THE QUEUE, CORRELATE THE CORRESPONDING REFERENCE MATRIX WITH THE TIME WARPED INPUT WAVEFORM S211
SELECT THE MAXIMUM CORRELATION VALUE FOR THE CURRENT SONG S213
IF CORRECTION PEAK IS ABOVE THE THRESHOLD, THAN DECLARE RECOGNITION & SAVE THE TIME, DATE, STATION AND SONG NUMBER S215
TO S300

F I G. 14